

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 39 seconds
(without alignments)

675.647 Million cell updates/sec

Title: US-09-806-536A-14

Perfect score: 1386

Sequence: 1 MAEAEPSPRRVGVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	613	44.2	279	2 T23765	hypothetical prote
2	584.5	42.2	319	1 A41770	pyrroline-5-carbox
3	502.5	36.3	273	2 T06477	probable pyrroline
4	492	35.5	274	1 S10186	pyrroline-5-carbox
5	480	34.6	270	2 G97299	pyrroline-5-carbox
6	475	34.3	276	1 JQ2334	pyrroline-5-carbox
7	469	33.8	266	2 AE1124	1-pyrroline-5-carb
8	458	33.0	266	2 AG1484	1-pyrroline-5-carb
9	455.5	32.9	270	2 AG1867	pyrroline-5-carbox
10	423.5	30.6	269	1 RDECC	pyrroline-5-carbox
11	423.5	30.6	269	2 D90683	pyrroline-5-carbox
12	423.5	30.6	269	2 H85533	pyrroline-5-carbox
13	420.5	30.3	284	2 T36286	pyrroline-5-carbox
14	419.5	30.3	269	2 A10549	pyrroline-5-carbox
15	412.5	29.8	267	2 S76767	probable pyrroline
16	395	28.5	265	2 H95107	pyrroline-5-carbox
17	390.5	28.2	263	2 D71281	probable pyrroline
18	387	27.9	261	2 JC2078	pyrroline-5-carbox
19	385	27.8	265	2 B97976	pyrroline-5-carbox
20	383.5	27.7	294	2 S72897	pyrroline-5-carbox
21	366	26.4	295	2 G70745	probable proc prot
22	362	26.1	299	2 T29226	hypothetical prote
23	360.5	26.0	277	2 H90194	hypothetical prote
24	354	25.5	270	2 E69682	pyrroline-5-carbox
25	354	25.5	279	2 G83837	pyrroline-5-carbox
26	349.5	25.2	264	2 C75385	pyrroline-5-carbox
27	349	25.2	262	2 D86860	pyrroline-5-carbox
28	345.5	24.9	273	2 JQ0418	pyrroline-5-carbox
29	344	24.8	282	2 T50305	Delta 1-pyrroline-

30	337.5	24.4	278	1 G69964	pyrroline-5-carbox
31	335	24.2	266	2 G72769	probable pyrroline
32	322	23.2	305	2 C82524	pyrroline-5-carbox
33	321	23.2	286	2 S25293	pyrroline-5-carbox
34	318.5	23.0	265	2 F70315	pyrroline carboxyl
35	312.5	22.5	267	2 G83760	pyrroline-5-carbox
36	310.5	22.4	311	2 S57863	pyrroline-5-carbox
37	305.5	22.0	320	2 JC4830	pyrroline-5-carbox
38	301.5	21.8	271	2 I64060	pyrroline-5-carbox
39	300.5	21.7	272	2 AH2847	pyrroline-5-carbox
40	300.5	21.7	274	2 G97624	delta 1-pyrroline-
41	298	21.5	272	2 D82321	pyrroline-5-carbox
42	297.5	21.5	255	2 E72360	pyrroline-5-carbox
43	290	20.9	260	2 E87310	pyrroline-5-carbox
44	283.5	20.5	300	2 AD3573	pyrroline-5-carbox
45	276	19.9	271	2 G89929	hypothetical prote

ALIGNMENTS

RESULT 1

T23765

hypothetical protein M153.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T23765

R:Matthews, P.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19796

A:Accession: T23765

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-279 <WIL>

A:Cross-references: EMBL:267995; PIDN:CAA91943.1; GSPDB:GN00028; CESP:M153.1

A:Experimental source: clone M153

C:Genetics:

A:Gene: CESP:M153.1

A:Map position: X

A:Introns: 23/1; 106/3; 177/3; 248/3

C:Superfamily: pyrroline-5-carboxylate reductase

Query Match	44.2%	Score	613;	DB	2;	Length	279;
Best Local Similarity	46.1%	Pred. No.	7.9e-41;				
Matches	123;	Conservative	56;	Mismatches	86;	Indels	2;
Gaps	1;						
Qy	10	RVGVFGAGRMAGATAOGLIRAGKVEAOHILASAP--TDRMLCHFQALGCRTHSHNQEVLQ	67				
Db	2	KIGFIGAGKMAQALARGLINSGRITADNIIASSPKRDEVFLDOCKALGLNTHDNAEVVQ	61				
Qy	68	SCLLVIFATKPHVLPVLAEVAPVVTTEHTILVSAAGVSLSTLEELLPPNTRVLRLVLPNL	127				
Db	62	KSDVVFVAVRPVHVSKVASEIAPALSREHLVSVIALGITIRNIESLPTTKSRVVRVMPNT	121				
Qy	128	PCVVOEGAIYVMARGHRVGSSETKLLOHLEACGCEEVPEAYVDIHTGLSGSGVAFVCAF	187				
Db	122	PSVVRACASAFAMSGACRDGDAETVKLLSTVGVFAVEVPEITHDPVTGLSGSGPSYMFVAV	181				
Qy	188	SEALAEAGVMKMGMPSSLAHRIAQTLLGTAKMLLHEQHQAQLRSDVCTPGGTTIYGLHA	247				
Db	182	IEGLADGGVKVGLPRDLALKLAAYTLGAAKMWLETGIHPAQLKDDVQSPAGSSVGMHK	241				
Qy	248	LEOGLRAATMSAVEAATCRAKELSRK	274				
Db	242	LESGLGLVLMDAVEAATNRSRATGDK	268				

RESULT 2

A41770

pyrroline-5-carboxylate reductase (EC 1.5.1.2) - human

N:Alternate names: P5C reductase

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AG1484
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC95647.1; PID:g16412843; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0414
C:Superfamily: pyrroline-5-carboxylate reductase

Query Match 33.0%; Score 458; DB 2; Length 266;
Best Local Similarity 37.5%; Pred. No. 1.1e-28;
Matches 98; Conservative 58; Mismatches 97; Indels 8; Gaps 2;
QY 9 RRVGFGAGRMAGATAOGLIRAGKVEAQHILASAPTDRLNCHFOAL-----GCRTHSNQ 63
DB 2 KKIGFIGAGMGAAMIRGLAKANLCKEEDILVCG---RDMEKLPKLTFFNGIQLTTDIA 58
QY 64 EVLOSCLLVIFATKPHVLPVLAELAVPVVTEHILVSVAGVSLSTLEELPPNTRVLRV 123
DB 59 QLAEQADIILSVKPYTIPPEILTAYKDKITPEKIVISVAGVTKIDLEELTSTETKIVRV 118
QY 124 LNLPCVQEGALVARGHVGSSSETKLLQHLLEACGRCEVEPYAYVDIHTGLSGGVAF 183
DB 119 MNTPALVCEAMSSISPSNVTAETAHLSIFSGRAEVAENLMDAVVGVSGSSPAY 178
QY 184 VCAFSEALAEAGVKMGPMSSLAHRTAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
DB 179 VYMFTEALADGAVLKGMPDRKAYKFAAQVLAQAQKWLTEGHPGKLDKMTVSPGGTTIE 238
QY 244 GLHALEOGLRAATMSAVEAA 264
DB 239 AVKSLEDTGRSSVISAVQAA 259

RESULT 9
AG1867
pyrroline-5-carboxylate reductase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AG1867
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG1867
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB72446.1; PID:g17129833; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0488
C:Superfamily: pyrroline-5-carboxylate reductase

Query Match 32.9%; Score 455.5; DB 2; Length 270;
Best Local Similarity 40.9%; Pred. No. 1.7e-28;
Matches 110; Conservative 43; Mismatches 109; Indels 7; Gaps 6;
QY 10 RVGFGAGRMAGATAOGLIRAGKVEAQHILASAP-TORNLCHFOALGCRTHSNQEVLOS 68
DB 4 KFLTGGVGMEALLSLIARTGIQYSEVIVSEPTARQAFLOQYHVGVTDTNSLVFTQ 63
QY 69 CLLVIF-ATKPHVLPVLAELAVPVVTEH--ILVSVAGVSLSTLEELPPNTRVLRVLP 125
DB 64 AQDVFLAVKQVFSATAQELADVTFTDHSPLVSVILAGVSLSLQEAFA-PQSPVIRAMP 122

QY 126 NLPCVQEGALVARGHVGSSSETKLLQHLLEACGRCEVEPYAYVDIHTGLSGGVAFVC 185
DB 123 NTPATVSGMTATCSGAYTAAQHQKLAQQIFSAVGEVVEVSEIMDAVTGLSGSPAYVA 182
QY 186 AFSEALAEAGVKMGPMSSLAHRTAAQTLLGTAKMLLHEGO-HPAQLRSDVCTPGGTTIY 244
DB 183 LLVEALADGGVASGLPRGIANQALQTVLGTAA-LLHESKLPALDKDRVTSPPGTTIAG 241
QY 245 LHAEQOGLRAATMSAVEAATCRAKELSR 273
DB 242 IAQLEKAGFRSALIEAVKAATWRSQELGK 270

RESULT 10
RDECC
pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 01-Mar-2002
C:Accession: A00385; B64767
R:Deutch, A.H.; Smith, C.J.; Rushlow, K.E.; Kretscher, P.J.
Nucleic Acids Res. 10, 7701-7714, 1982
A:Title: Escherichia coli delta(1)-pyrroline-5-carboxylate reductase: gene sequence,
A:Reference number: A00385; MUID:83116986; PMID:6296787
A:Accession: A00385
A:Molecule type: DNA
A:Residues: 1-269 <DEU>
A:Cross-references: GB:J01655; NID:g147358; PIDN:AAA86433.1; PID:g147359
A:Note: parts of this sequence, including the amino and carboxyl ends of the mature p
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64767
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-269 <BLAT>
A:Cross-references: GB:AE000145; GB:U00096; NID:g1786580; PIDN:AAC73489.1; PID:g17865
A:Experimental source: strain K-12, substrain MGL655
C:Genetics:
A:Gene: proc
A:Map position: 9 min
C:Function:
A:Description: catalyzes reduction of pyrroline-5-carboxylate to proline
A:Pathway: proline biosynthesis
A:Note: third enzyme in the proline biosynthetic pathway
C:Superfamily: pyrroline-5-carboxylate reductase
C:Keywords: oxidoreductase; proline biosynthesis

Query Match 30.6%; Score 423.5; DB 1; Length 269;
Best Local Similarity 34.6%; Pred. No. 5.6e-26;
Matches 92; Conservative 56; Mismatches 117; Indels 1; Gaps 1;
QY 9 RRVGFGAGRMAGATAOGLIRAGKVEAQHILASAPT-DRNLCHFOALGCRTHSNQEVLIQ 67
DB 3 KKIGFTCCGMGKAILGLGLASQVLPQGIWVYTPSPDKVAALHDFGGINAAEQEVAQ 62
QY 68 SCLLVIFATKPHVLPVLAELAVPVVTEHILVSVAGVSLSTLEELPPNTRVLRVLPML 127
DB 63 IADITFAAVKPGIMTKVLSITSSLNKDSLVVIAAGVTLQDLARALGHDRIIRAMPNT 122
QY 128 PCVVOEGALVARGHVGSSSETKLLQHLLEACGRCEVEPYAYVDIHTGLSGGVAFVCAF 187
DB 123 PALVNAAGMTSVITPNALVTPEDTADVLNIFRCFGEAEVTAEPMIHPVVGSGSSPAYVPMF 182
QY 188 SEALAEAGVKMGPMSSLAHRTAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
DB 183 IEAMADAVALGMPRAQAYKFAAQAVMGSAKMWLETGEHPGALKDMVCSPPGTTIEAVRV 242
QY 248 LEQOGLRAATMSAVEAATCRAKELSR 273
DB 243 LEEKFRAAVTEAMTKCMKSEKLSK 268

RESULT 11
 D90683
 pyrroline-5-carboxylate reductase [imported] - Escherichia coli (strain O157:H7, substra
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: D90683
 R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: D90683
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-269 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA33859.1; PID:g13359893; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs0437
 C:Superfamily: pyrroline-5-carboxylate reductase
 Query Match 30.6%; Score 423.5; DB 2; Length 269;
 Best Local Similarity 34.6%; Pred. No. 5.6e-26;
 Matches 92; Conservative 56; Mismatches 117; Indels 1; Gaps 1;
 QY 9 RRVGFGVAGRMAGATAOGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTHHNOEVLO 67
 DB 3 KKGIFGCGNGMKAILGGLIASGOVLPGQIWYTPSPDKVAALHDKFGINAAESAQVAV 62
 QY 68 SCLLVIFATKPHVLPVLAEPVVTTEHILSVSVAAGVSLTLEELPPNTRVRLVLPNL 127
 DB 63 IADIIFAAVKPGIMIKVLSEITSSLNKDSLVVSIAGVTLDLQALRALGHDKRIIRAMPNT 122
 QY 128 PCVVOEGAIYVARGHVGSSSETKLLQHLLEACRCCEVPAYVDIHTGLSGSGVAFVCF 187
 DB 123 PALVNAGMTSVTPNALVTPEADVLNIFRCFGEAEVIAEPMIHPVVGSSSPAYVFMF 182
 QY 188 SEALAEAGVAKMGMPSSLAHRIAOTLLGTAKMLLHEGHPAQLRSDVCTPGGTTIYGLHA 247
 DB 183 IEAMADAALVGGMPRAQAYKFAAQVAMGSAKVLETGEHPGALKDMVCSGGTTIEAVRV 242
 QY 248 LEQGLRAATMSAVEAATCRAKELSR 273
 DB 243 LEEKGFRAAVIEAMTKMEKSEKLSK 268
 RESULT 12
 H85533
 pyrroline-5-carboxylate reductase [imported] - Escherichia coli (strain O157:H7, substra
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85533
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85533
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-269 <STO>
 A:Cross-references: GB:AE005174; NID:g12513226; PIDN:AAG54732.1; GSPDB:GN00145; UWGP:204
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: proc
 C:Superfamily: pyrroline-5-carboxylate reductase
 Query Match 30.6%; Score 423.5; DB 2; Length 269;
 Best Local Similarity 34.6%; Pred. No. 5.6e-26;
 Matches 92; Conservative 56; Mismatches 117; Indels 1; Gaps 1;
 QY 9 RRVGFGVAGRMAGATAOGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTHHNOEVLO 67
 DB 3 KKGIFGCGNGMKAILGGLIASGOVLPGQIWYTPSPDKVAALHDKFGINAAESAQVAV 62
 QY 68 SCLLVIFATKPHVLPVLAEPVVTTEHILSVSVAAGVSLTLEELPPNTRVRLVLPNL 127
 DB 63 IADIIFAAVKPGIMIKVLSEITSSLNKDSLVVSIAGVTLDLQALRALGHDKRIIRAMPNT 122
 QY 128 PCVVOEGAIYVARGHVGSSSETKLLQHLLEACRCCEVPAYVDIHTGLSGSGVAFVCF 187
 DB 123 PALVNAGMTSVTPNALVTPEADVLNIFRCFGEAEVIAEPMIHPVVGSSSPAYVFMF 182
 QY 188 SEALAEAGVAKMGMPSSLAHRIAOTLLGTAKMLLHEGHPAQLRSDVCTPGGTTIYGLHA 247
 DB 183 IEAMADAALVGGMPRAQAYKFAAQVAMGSAKVLETGEHPGALKDMVCSGGTTIEAVRV 242
 QY 248 LEQGLRAATMSAVEAATCRAKELSR 273
 DB 243 LEEKGFRAAVIEAMTKMEKSEKLSK 268
 RESULT 13
 T36286
 pyrroline-5-carboxylate reductase - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T36286
 R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21603
 A:Accession: T36286
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-284 <SEE>
 A:Cross-references: EMBL:AL049819; PIDN:CA842663.1; GSPDB:GN00070; SCOEDB:SCE7.04c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: proC; SCOEDB:SCE7.04c
 C:Superfamily: pyrroline-5-carboxylate reductase
 Query Match 30.3%; Score 420.5; DB 2; Length 284;
 Best Local Similarity 35.4%; Pred. No. 1e-25;
 Matches 95; Conservative 53; Mismatches 119; Indels 1; Gaps 1;
 QY 5 EPSPRVGVGAGRMAGATAOGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTHHNOE 64
 DB 13 EPMTQKVAVLGTGKIGEALLSGMIGAGWAPAD-LLVTAARRERADELRARHGVTPTVNAE 71
 QY 65 VLQSCLLVIFATKPHVLPVLAEPVVTTEHILSVSVAAGVSLTLEELPPNTRVRLV 124
 DB 72 AAKAADTLITVKPQDMGTLDELAPHVPADRLVISGAAGVPTSFEEELAPGTPVVRVM 131
 QY 125 PNLPCVVOEGAIYVARGHVGSSSETKLLQHLLEACRCCEVPAYVDIHTGLSGSGVAFV 184
 DB 132 TNPALVDEAMSVISAGTHATAHLTHTTEIFCAVGKTLRVPESSQODACTALSGSGPAYF 191
 QY 185 CAFSEALAEAGVAKMGMPSSLAHRIAOTLLGTAKMLLHEGHPAQLRSDVCTPGGTTIY 244
 DB 192 FYLVNMTDAGILLGLPRDKAHDLLIVQSAIGAAMKLRDSGEHPVKLRNVTSPAGTTINA 251
 QY 245 LHALEQGLRAATMSAVEAATCRAKELS 272
 DB 252 IRELENHGVRAALIAALEAARDRSRELA 279
 RESULT 14
 AI0549
 pyrroline-5-carboxylate reductase [imported] - Salmonella enterica subsp. enterica se
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AI0549
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AI0549
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-269 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD08841.1; PID:g16501655; GSPDB:GN00176
C:Genetics:
A:Gene: STY0419
C:Superfamily: pyrroline-5-carboxylate reductase

Query Match 30.3%; Score 419.5; DB 2; Length 269;
Best Local Similarity 34.6%; Pred. No. 1.2e-25;
Matches 92; Conservative 57; Mismatches 116; Indels 1; Gaps 1;

QY 9 RVGVFGAGRMAGAIAGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTHSNQEVQLQ 67
Db 3 KKIGFIGCGNGKAILGGLIATASGOVLPQIWWYTPSPDKVAALHDQYGINAAQSAQEVQ 62

QY 68 SCLLVIFATKPHVLPVLAELVAVPVVTTTEHILSVAGVSLSTLEELLPPNTRVLRVLPNL 127
Db 63 VADIVFGAVKPGIMVKVLSLSEISSLNKDSLWSTAGVTLDLQARALGHRKTIIRAMPNT 122

QY 128 PCVVOEGAIVMARGHVSSETKLLQHLLEACGCEVPEAYVDIHTGLSGSGVAFVCAF 187
Db 123 PSLVNAAGMTVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHPVVGVSQSPAYVFMF 182

QY 188 SEALAEGAVKMGMPSSLAHRIAQTLLGTAKMLLHEGHPAQLRSDVCTPGGTTIYGLHA 247
Db 183 IEAMDAAVLGGMPRAQYKFAAQAVMGSAKIVLETKHPELKDWCSPCGTTIEAVRV 242

QY 248 LEQGLRAATNSAVEAATCRAKELSR 273
Db 243 LEERGFRAAVTEAMTKMEKSEALSK 268

RESULT 15
S76767
probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - *Synechocystis* sp. (strain PCC
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76767
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76767
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-267 <KAN>
A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAAL8679.1; PID:g165376
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG
C:Superfamily: pyrroline-5-carboxylate reductase
C:Keywords: oxidoreductase; proline biosynthesis

Query Match 29.8%; Score 412.5; DB 2; Length 267;
Best Local Similarity 36.2%; Pred. No. 4.1e-25;
Matches 98; Conservative 50; Mismatches 110; Indels 13; Gaps 4;

QY 10 RVGVFGAGRMAGAIAGLIRAGKVEAQHILASAPTDRNLCHFQ-ALGCRTHSNQEVQLQ 68
Db 4 QLGIIGGVMAEAILARLIAEKTYAPEIIVGEPHGARRDYLQKTYQVRVSPDNQEAANV 63

QY 69 CLLVIFATKPHVLPVLAELVAVPVVTTTEHILSVAGVSLSTLEELLPPNTRVLRV 123

Db 64 SEVLLAVKPOVLDRLVSLASLAGANRPLV------ISILAGVSLQRIQKGFDPHA-IIRA 116
QY 124 LPNLFCVVOEGAIVMARGHVSSETKLLQHLLEACGCEVPEAYVDIHTGLSGSGVAF 183
Db 117 MPNTPATYTGAGMTAIAANKWVEPDOLAKAKAIFSAVGNVVEVPENLMDAVTGVSGSPAY 176
QY 184 VCAFSEALAEGAVKMGMPSSLAHRIAQTLLGTAKMLLHEGHPAQLRSDVCTPGGTTIY 243
Db 177 VALMTEALADGGVLAGLPRAIAQKIALQTLVLTAEILKETEHHPAQIKDKVTSFGGTTIA 236
QY 244 GLHALEQGLRAATNSAVEAATCRAKELSRK 274
Db 237 GVAVLEKMGFRSAIIEAVRAAYRRSQELGKK 267

Search completed: July 30, 2003, 15:28:05
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:57 ; Search time 22 Seconds
(without alignments)
1479.099 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRVGVGAGRMA.....AATMSAVEATCRANKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1386	100.0	274	15 US-10-161-418A-11	Sequence 11, Appl
2	1386	100.0	274	15 US-10-161-418A-13	Sequence 13, Appl
3	584.5	42.2	319	15 US-10-161-418A-10	Sequence 10, Appl
4	584.5	42.2	319	15 US-10-161-418A-12	Sequence 12, Appl
5	560.5	40.4	315	9 US-09-912-717-3	Sequence 3, Appl
6	537.5	38.8	314	9 US-09-912-717-1	Sequence 1, Appl
7	477	34.4	319	15 US-10-161-418A-14	Sequence 14, Appl
8	467.5	33.7	255	10 US-09-925-300-1218	Sequence 1218, Ap
9	408.5	29.5	269	15 US-10-156-761-12258	Sequence 12258, A
10	401	28.9	284	15 US-10-128-714-3252	Sequence 3252, Ap
11	401	28.9	284	15 US-10-128-714-8252	Sequence 8252, Ap
12	366	26.4	295	10 US-09-712-363-174	Sequence 174, App
13	362.5	26.2	270	10 US-09-738-626-3960	Sequence 3960, Ap
14	116	8.4	144	9 US-09-939-980-445	Sequence 445, App
15	114	8.2	299	9 US-09-912-020-256	Sequence 256, App

16	95.5	6.9	549	15	US-10-156-761-14029	Sequence 14029, A
17	95	6.9	807	11	US-09-930-020A-2	Sequence 2, Appli
18	93.5	6.7	289	15	US-10-156-761-9583	Sequence 9583, Ap
19	93	6.7	2436	9	US-09-795-693-8	Sequence 8, Appli
20	93	6.7	2436	15	US-10-156-761-239-8	Sequence 8, Appli
21	93	6.7	2436	15	US-10-156-761-239-8	Sequence 8, Appli
22	92	6.6	408	10	US-09-712-363-273	Sequence 273, App
23	91.5	6.6	452	15	US-10-156-761-12493	Sequence 12493, A
24	89.5	6.5	323	15	US-10-156-761-9372	Sequence 9372, Ap
25	89.5	6.5	336	15	US-10-156-761-10217	Sequence 10217, A
26	89.5	6.5	488	15	US-10-156-761-7949	Sequence 7949, Ap
27	89.5	6.5	858	9	US-09-815-242-11396	Sequence 11396, A
28	89	6.4	447	15	US-10-156-761-9880	Sequence 9880, Ap
29	87.5	6.3	1024	15	US-10-211-962-44	Sequence 44, Appl
30	87.5	6.3	6145	15	US-10-156-761-7962	Sequence 7962, Ap
31	87	6.3	426	15	US-10-156-761-8581	Sequence 8581, Ap
32	87	6.3	949	9	US-09-841-835-10	Sequence 10, Appl
33	87	6.3	1327	9	US-09-841-835-2	Sequence 2, Appli
34	87	6.3	1327	11	US-09-972-115A-8	Sequence 8, Appli
35	86.5	6.2	340	9	US-09-815-242-10852	Sequence 10852, A
36	86.5	6.2	342	9	US-09-815-242-4924	Sequence 4924, Ap
37	86.5	6.2	1835	15	US-10-156-761-7963	Sequence 7963, Ap
38	86.5	6.2	4809	15	US-10-156-761-9090	Sequence 9090, Ap
39	86	6.2	492	9	US-09-815-242-11079	Sequence 11079, A
40	86	6.2	499	15	US-10-156-761-9271	Sequence 9271, Ap
41	85.5	6.2	342	15	US-10-156-761-14291	Sequence 14291, A
42	85.5	6.2	748	9	US-09-815-242-12792	Sequence 12792, A
43	85.5	6.2	792	9	US-09-815-242-12327	Sequence 12327, A
44	85.5	6.2	1687	14	US-10-094-679-3	Sequence 3, Appli
45	85	6.1	1074	9	US-09-509-196A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-161-418A-11
; Sequence 11, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT APPLICATION NUMBER: US/10/161,418A
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-418A-11

Query Match	100.0%	Score	1386	DB	15	Length	274
Best Local Similarity	100.0%	Pred. No.	1.4e-134				
Matches	274	Conservative	0	Mismatches	0	Indels	0
						Caps	0
Qy	1	MAAAEPSPRVGVGAGRMAAGTAQGLIRAGKVEAOHILASAPTDRLNCHFGALGCRTHH	60				
Db	1	MAAAEPSPRVGVGAGRMAAGTAQGLIRAGKVEAOHILASAPTDRLNCHFGALGCRTHH	60				
Qy	61	SNOEVLQSCLLVIFATKPHVLPVLAVEAPVVTTEHILVSVAAAGVSLSTLEELLPNTRV	120				
Db	61	SNOEVLQSCLLVIFATKPHVLPVLAVEAPVVTTEHILVSVAAAGVSLSTLEELLPNTRV	120				
Qy	121	LRVLNPLPCVVQEGCAIWMARGHVGSGSETKLLQHLLEACGRCBEVPEAYVDIHTGLSGG	180				
Db	121	LRVLNPLPCVVQEGCAIWMARGHVGSGSETKLLQHLLEACGRCBEVPEAYVDIHTGLSGG	180				

```
QY 181 VAFVCAFSEALAEAGVKGMPSSLAHRIAATLLGTAKMLLHEGHPAQLRSDVCTPGGT 240
Db 181 VAFVCAFSEALAEAGVKGMPSSLAHRIAATLLGTAKMLLHEGHPAQLRSDVCTPGGT 240
QY 241 TIYGLHALEOQGLRAATMSAVEAATCRAKELSRK 274
Db 241 TIYGLHALEOQGLRAATMSAVEAATCRAKELSRK 274

RESULT 2
US-10-161-418A-13
; Sequence 13, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT APPLICATION NUMBER: US/10/161,418A
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-418A-13
```

```
Query Match 100.0%; Score 1386; DB 15; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.4e-134;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAEPSPRVGFAGRMAGATAOGLIRAGKVEAOHILASAPTDNRNLCHFQAL 60
Db 1 MAAAEPSPRVGFAGRMAGATAOGLIRAGKVEAOHILASAPTDNRNLCHFQAL 60
QY 61 SNOEVLQSCLLVIFATKPHVLPVLAEPVVTTEHILVSVAAAGVSLSTLEELLPPNTRV 120
Db 61 SNOEVLQSCLLVIFATKPHVLPVLAEPVVTTEHILVSVAAAGVSLSTLEELLPPNTRV 120
QY 121 LRVLPNLCVVOEGAIVMARGHRVGSSETKLLOHLLACGRCCEEPVAYVDIHTGLSGSG 180
Db 121 LRVLPNLCVVOEGAIVMARGHRVGSSETKLLOHLLACGRCCEEPVAYVDIHTGLSGSG 180
QY 181 VAFVCAFSEALAEAGVKGMPSSLAHRIAATLLGTAKMLLHEGHPAQLRSDVCTPGGT 240
Db 181 VAFVCAFSEALAEAGVKGMPSSLAHRIAATLLGTAKMLLHEGHPAQLRSDVCTPGGT 240
QY 241 TIYGLHALEOQGLRAATMSAVEAATCRAKELSRK 274
Db 241 TIYGLHALEOQGLRAATMSAVEAATCRAKELSRK 274
```

```
RESULT 3
US-10-161-418A-10
; Sequence 10, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT APPLICATION NUMBER: US/10/161,418A
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
```

```
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-418A-10

Query Match 42.2%; Score 584.5; DB 15; Length 319;
Best Local Similarity 46.4%; Pred. No. 6.6e-52;
Matches 124; Conservative 47; Mismatches 89; Indels 7; Gaps 3;

QY 11 VGFVGAGRMAGATAOGLIRAGKVEAOHILASAPTDNRNLCHFQAL---GCRTHSNOEVLQ 67
Db 3 VGFIGAQQLAFALAKGFTAAGVLAHAKIMASSP-DMDLATVSAIRKMGVKLTTPHNKETVQ 61
QY 68 SCLLVIFATKPHVLPVLAEPVVTTEHILVSVAAAGVSLSTLEELLP---PNTRVLRVL 124
Db 62 HSDVLFVAVKPHIIPFLDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRAPVIRCM 121
QY 125 PNLPCVVOEGAIVMARGHRVGSSETKLLOHLLACGRCCEEPVAYVDIHTGLSGGVAFV 184
Db 122 TNTPVVREGATVYATGTHAQVEDGRLEQLLSTVGCTEVEEDLIDAVTGLSGGPAYA 181
QY 185 CAFSEALAEAGVKGMPSSLAHRIAATLLGTAKMLLHEGHPAQLRSDVCTPGGTTIYG 244
Db 182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEHPGQOLKDNVSSPGGATIHA 241
QY 245 LHAEQGLRAATMSAVEAATCRAKEL 271
Db 242 LHVLESGGFRLLINAVEASCIRTREL 268
```

```
RESULT 4
US-10-161-418A-12
; Sequence 12, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT APPLICATION NUMBER: US/10/161,418A
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-418A-12
```

```
Query Match 42.2%; Score 584.5; DB 15; Length 319;
Best Local Similarity 46.4%; Pred. No. 6.6e-52;
Matches 124; Conservative 47; Mismatches 89; Indels 7; Gaps 3;

QY 11 VGFVGAGRMAGATAOGLIRAGKVEAOHILASAPTDNRNLCHFQAL---GCRTHSNOEVLQ 67
Db 3 VGFIGAQQLAFALAKGFTAAGVLAHAKIMASSP-DMDLATVSAIRKMGVKLTTPHNKETVQ 61
QY 68 SCLLVIFATKPHVLPVLAEPVVTTEHILVSVAAAGVSLSTLEELLP---PNTRVLRVL 124
Db 62 HSDVLFVAVKPHIIPFLDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRAPVIRCM 121
QY 125 PNLPCVVOEGAIVMARGHRVGSSETKLLOHLLACGRCCEEPVAYVDIHTGLSGGVAFV 184
Db 122 TNTPVVREGATVYATGTHAQVEDGRLEQLLSTVGCTEVEEDLIDAVTGLSGGPAYA 181
QY 185 CAFSEALAEAGVKGMPSSLAHRIAATLLGTAKMLLHEGHPAQLRSDVCTPGGTTIYG 244
Db 182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEHPGQOLKDNVSSPGGATIHA 241
QY 245 LHAEQGLRAATMSAVEAATCRAKEL 271
```


Db 177 -AFMALDADGGYKMLPRRLAQLGAQALLGAARKMLDSEHPQLKDNVCSPGATIHA 235

Qy 245 LHALPQGGILRAATMSAVEAATCRAKEL 271

Db 236 LHFESGGFRSLINAVEASCIITREL 262

```

RESULT 7
US-10-161-418A-14
; Sequence 14, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT APPLICATION NUMBER: US/10/161,418A
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-418A-14

```

Query Match	34.4%	Score 477:	DB 15:	Length 319:
Best Local Similarity	41.0%:	Pred. No. 7.9e-41:		
Matches 110:	Conservative 47:	Mismatches 101:	Indels 10:	Gaps 4:
QY	11	VGFVGAGRMA-----GATAOGLIRAGKVEAQHILASAPTDNLCHEQALGCRTHHSNQEVFL	66	
Db	3	VGFIGAGQLAMLRGA--SRQIPVGSODNSQLPRNEPA--HGVRAQEDCCEDPTQOQGDG	59	
QY	67	QSCLLVIFATKPKHPLAVLAEPVPTTEHILVSVGAAGVSLSTLBELL---PPNTRVLRLV	123	
Db	60	EAQRPPVSGCGAHIIPIFVIEIGADVOARHIVVSCAAGVTISVSEKLMETQPAKPVIRC	119	
QY	124	LPNLPCVVQEGAIYMAGRHVGSSETKLLQHLLEACGRCGEEVPAYVDIHGLSGSGVAF	183	
Db	120	MTNPVWVQEGATVYATGTHTALVEDGQLLEQLMSSVGFCTEVEDLIDAVTGLSGSRPAY	179	
QY	184	VCATSEALACAVKMGKMPSSLAHRIAAQTLLGTAKMLLHGQHQAQLRSDVCTPGGTIY	243	
Db	180	AFMALDALADGGVKNMGPRLRLATOLGAQALLGAAKMLLDSEQHPCQLKDNVCSFGGATIH	239	
QY	244	GLHALEQGGLRAATMSAVEZATCRAKEL	271	
Db	240	ALHFLESGGFSSLLINAVEASLRTRTEL	267	

```

RESULT 8
US-09-925-300-1218
; Sequence 1218, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925, 300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124, 270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1218

```

```

; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1218

```

Query Match	33.7%	Score 467.5;	DB 10;	Length 255;
Best Local Similarity	48.5%;	Pred. No. 5.5e-40;		
Matches	95;	Conservative 34;	Mismatches 64;	Indels 3; Gaps 1;
QY	79	HVLPAVLAEVAPVVTTEHILVSVAAAGVSLSTLELLP--PNTRLVRLVLPNLPCVVOEGA	135	
Db	9	HIIPILDEICADIEDRHVVSCAAGVYIISIIKKLSAFAPRVIRCMINTPVPVVRGA	68	
QY	136	IYVAGRHVGSSEYKLLQHLGLEACGRCEPEAYVDIHTGLSGSGVAFVCAFEALAEGA	195	
Db	69	TYVATGTTHAQVEDGRLEMEQLLSVYGFCTEVEEDIDAVTGLSGSGPAPAVFTALDALADGG	128	
QY	196	VKMGMPSLLAHRIAATLLGTAKMLLHEGOHPAQLRSDVCTPGGTTIVGLHALEGGGLRA	255	
Db	129	VKMGLPRLRAVLRCGAQLLGAAKMLLHSEQHPQQLKDNVSSPGGATIIHALHVLESGGFRS	188	
QY	256	ATMSAVEAATCRAKEL	271	
Db	189	LLINAVEASCIRTREL	204	

```

RESULT 9
US-10-156-761-12258
; Sequence 12258, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIORITY APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12258
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12258

```

[illegible]


```

; LENGTH: 270
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3960

Query Match      26.2%   Score 362.5; DB 10; Length 270;
Best Local Similarity 30.6%; Pred. No. 3.9e-29;
Matches 83; Conservative 63; Mismatches 114; Indels 11; Gaps 4;

QY 11 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTDRLNLCHFQAL---GCRTHHSNOEVL 66
Db 4 IAVIGGQIGEALVSGLI-AANNPNQIRV---TNRSEERGQELDRYGLINMTDNSQAA 59
QY 67 QSCLLVIFATKPHVLPAVLAEVAPV---TTEHILSVSAAGVSLSTELELLPENTRVLRV 123
Db 60 DEADVVFLCVKPKFIEVLEITGTLDNNSAQVSVYMAAGISIAAMEESASAGLPVVRV 119
QY 124 LPNLPCVVOBGAIVMARGHVGSEYTKLLOHLLAEACRCEVEPEAYVDHTGLSGSGVAF 183
Db 120 MPNTPLMLVGKGMTVTKGRYVDRAEQLEQVKDILLSTVGDVLEVAESDIDAVTAMSGSPAY 179
QY 184 VCAFEALBAGYKMGMPSSLAHRIAQAOTLLGTAKMLLHEGQHPAQLRSRDVCTPGGTTIV 243
Db 180 LFLVTEALIEAGVNLGLIPRATAKKLAVASFEAGATWMKETGKPESELACGVSSPACTTVA 239
QY 244 GLHALFQGLRAATMSAVEANATCRAKELSRK 274
Db 240 AIRELEESGTRGAFYRAAQACADRSPELGKR 270

RESULT 14
US-09-939-980-445
; Sequence 445, Application US/09939980
; Patent No. US20020082234A1
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; Burnham, Martin
; Hodgson, John
; Knowles, David
; Lonetto, Michael
; Nicholas, Michael
; Pratt, Julie
; Reichard, Richard
; Rosenberg, Martin
; Ward, Judith
; TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
; Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmli, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-320-4479

```

```

; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 445:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 445:
US-09-939-980-445
```

```

Query Match      8.4%; Score 116; DB 9; Length 144;
Best Local Similarity 26.8%; Pred. No. 0.0004;
Matches 30; Conservative 24; Mismatches 56; Indels 2; Gaps 2;

QY 16 AGRMAGATAOGLIRACKVEAQHI-LASAPTRDNLCHF-QALGCRTHSNQEVLOSCLLVI 73
Db 12 AGNMAQAIFTGINSNDANDIYLTNKSNEQALKAFKLGVNSYDDATLLKDADYVF 71
QY 74 FATKPHVLPVLAEAVPVVTEHILVSVAAAGVSLSTLELLPPNTRVLRVLP 125
Db 72 LGTKPHDFDALATRIKPHITKDXCFNSIWAGIPTDIYXQLECONPXARIMP 123
```

```

RESULT 15
US-09-912-020-256
; Sequence 256, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Forsyth, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ELITRA.001DV1
; CURRENT APPLICATION NUMBER: US/09/912,020
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/492,709
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 256
; LENGTH: 299
; TYPE: PRT
; ORGANISM: E. Coli
US-09-912-020-256
```

```

Query Match      8.2%; Score 114; DB 9; Length 299;
Best Local Similarity 21.7%; Pred. No. 0.0018;
Matches 69; Conservative 46; Mismatches 119; Indels 84; Gaps 16;

QY 10 RVGFVAGRMAGAIAGGLIRACKVEAQHILASAPTRDN---LCHFQALGCRTHSNQEV 66
Db 7 KVGFIGLGTMGKPMKSNLLKAG-----YSLVADRNPETADVIAGAGETASTAKAIA 59
QY 67 QSC--LLVIFATKPHVLPVLAEAVPVVTEHILVSVAAAGVSLSTLELLPPNTRVLRV 124
Db 60 EQCDVITITMLPNSPHKEVALGE-----NGIEGAKPGTVLIDMSSIAPLASREI--- 109
QY 125 PNLPCVVOEGATVMARG-----RHVGSSETKLLQHLLEA-CGRCEEVPEAYVD----- 171
Db 110 -----SEALKAKGIDMLDAPVSGGEPKAI DGTLSVMVGGDKAIFDKYIDLMKAMAG 160
QY 172 --IHTGLSGSG-----VAF-VCAFSALAEAGVAKMGMPSSLAHR-----IAAOTL 213
```

```

Db 161 SVVHTGETGAGNVTKLANQVIVALNIAAMSEALT-ATKAGVNPDLVYQAIRGGLAGSTV 219
QY 214 L-GTAKMILLHEGQHPA-----QLRSDVCTPGG-----TTIYGLHALEQGGGLRA 255
Db 220 LDAKAPVMVMDRNFKPGPRIDLHIKDLANALDTSHGVCQAQLPLTAAVNEMMQALRADGLGT 279
QY 256 ATMSAVEAATCRAKELSR 273
Db 280 ADHSAL---ACYYEKLAK 294

Search completed: July 30, 2003, 15:29:11
Job time : 23 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 85 Seconds
(without alignments)
511.660 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSRRVGVGAGRMA.....AATMSAVEATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	1386	100.0	274	21	Human OXRE-14. Ho
2	1386	100.0	274	22	Human protein sequ
3	1386	100.0	274	24	P5CR related polyp
4	1386	100.0	274	24	P5CR related polyp
5	1383	99.8	274	22	Human dihydroxyro
6	590.5	42.6	343	23	Human ovarian anti
7	584.5	42.2	319	22	Human P5CR protein
8	584.5	42.2	319	22	P5CR related polyp
9	584.5	42.2	319	24	P5CR related polyp

10	577.5	41.7	320	22	AAE12784	Human delta 1-pyrr
11	560.5	40.4	273	22	ABB61855	Drosophila melanog
12	560.5	40.4	315	22	AAE12785	Human delta 1-pyrr
13	560.5	40.4	320	22	AAE12785	Human delta 1-pyrr
14	537.5	38.8	314	21	AAAB20584	Human delta 1-pyrr
15	537.5	38.8	314	21	AAAU99322	Human delta 1-pyrr
16	505.5	36.5	234	24	ABR41145	Human DTHP enzyme
17	505.5	36.5	280	22	ABR71145	Drosophila melanog
18	477	34.4	319	24	AAAG79746	P5CR related polyp
19	475	34.3	276	21	AAAG05978	Arabidopsis thalia
20	469	33.8	266	23	ABBA7427	Listeria monocytog
21	467.5	33.7	255	21	AAAB56640	Human prostate can
22	436	31.5	256	21	AAAG06979	Arabidopsis thalia
23	423.5	30.6	269	22	AAAG64106	Escherichia coli p
24	401	28.9	284	24	ABJ25594	Aspergillus fumiga
25	401	28.9	284	24	ABJ25594	Aspergillus fumiga
26	395	28.5	265	24	AHU01336	S. pneumoniae type
27	392.5	28.3	268	22	AAU050524	Propionibacterium
28	388	28.0	212	22	AAAB93027	Human protein sequ
29	366	26.4	295	22	AAAG81123	Mycobacterium tube
30	362.5	26.2	270	22	AAAG90206	C glutamicum prote
31	362.5	26.2	270	22	AAAG90206	Corynebacterium gl
32	360.5	26.0	274	23	ABP73860	Candida albicans e
33	350.5	25.3	326	23	ABP65790	Lactobacterium lo
34	349	25.2	262	23	ABB55233	Streptococcus lactis
35	344.5	24.9	256	23	ABP25621	Streptococcus poly
36	343	24.7	90	22	AAU23472	Novel human enzyme
37	325	23.4	83	22	AAU23493	Novel human enzyme
38	316.5	22.8	256	23	ABP25622	Streptococcus poly
39	309	22.3	151	21	AAAG06980	Arabidopsis thalia
40	269	19.4	271	22	AAAG82031	S. epidermidis ope
41	269	19.4	282	23	ABP38310	Staphylococcus epi
42	251.5	18.1	263	23	AAAG66078	N. meningitidis pr
43	245.5	17.7	108	22	ABG20692	Novel human diagno
44	244.5	17.6	253	24	ABF80908	N. gonorrhoeae ami
45	202.5	14.6	279	23	ABB49027	Listeria monocytog

ALIGNMENTS

RESULT 1
AA92517
ID AAY92517 standard; Protein: 274 AA.

XX AAY92517;

XX 10-AUG-2000 (first entry)

DT Human OXRE-14.

DE OXRE-14; oxidoreductase; pyrroline-5-carboxylate-reductase;

DE antiproliferative; anticancer; hepatotropic; antiviral; antiaesthatic;

KW neotropic; neuroprotectant; antiparkinsonian's; antisclerotic;

KW anxiolytic; antischizophrenic; anti-inflammatory; antiarthritic.

OS Homo sapiens.

XX Key

XX Region

XX Location/Qualifiers

FT 9..256

FT /label= signature_sequence

FT /note= "for delta 1-pyrroline-5-carboxylate reductase"

FT 7

FT Modified-site

FT /note= "potential phosphorylation site"

FT 44

FT Modified-site

FT /note= "potential phosphorylation site"

FT 61

FT Modified-site

FT /note= "potential phosphorylation site"

FT 108

FT Modified-site

FT /note= "potential phosphorylation site"

FT 109

FT Modified-site

FT /note= "potential phosphorylation site"

FT 109

FT Modified-site

FT /note= "potential phosphorylation site"

FT 216

FT Modified-site

FT Modified-site /note= "potential phosphorylation site"
 FT 259
 FT Modified-site /note= "potential phosphorylation site"
 FT 265
 FT Modified-site /note= "potential phosphorylation site"
 XX
 PN WO200020604-A2.
 XX
 XX 13-APR-2000.
 XX
 XX 06-OCT-1999; 99WO-US23434.
 XX
 XX 06-OCT-1998; 98US-0172227.
 PR 02-DEC-1998; 98US-0155202.
 PR 10-MAR-1999; 99US-0123911.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Lal P, Guegler KJ, Gorgone GA, Corley NC, Baughn MR, Tang YT;
 PI Hillman JL, Bandman O, Azimzai Y, Au-Young J, Yue H, Lu DAM;
 PI Yang J;
 XX
 DR WPI; 2000-303785/26.
 DR N-PSDB; AAA09388.
 XX
 PT Purified polypeptide for treating or preventing disorders associated
 PT with decreased expression or activity of oxidoreductase molecules
 XX
 XX Claim 1; Page 84; 97pp; English.
 XX
 PS This OXRE-14 has identity with pyrroline-5-carboxylate-reductase.
 CC The polypeptides are useful for treating or preventing a disorder
 CC associated with decreased expression or activity of OXRE. Antagonists of
 CC OXRE are useful for treating or preventing a disorder associated with
 CC increased expression or activity of OXRE. The disorders include cell
 CC proliferative disorders (cirrhosis, hepatitis), cancer (leukemia, and
 CC melanoma), hypopituitarism and hyperpituitarism, hypothyroidism and
 CC hyperthyroidism, metabolic disorders (Addison's disease, cystic
 CC fibrosis), reproductive disorders (infertility, ovulatory defects),
 CC neurological disorders (Alzheimer's disease, Parkinson's disease,
 CC multiple sclerosis), mental disorders (anxiety, schizophrenia),
 CC autoimmune/inflammatory disorders (acquired immunodeficiency syndrome
 CC (AIDS), asthma, osteoarthritis), and viral infections. The
 CC polynucleotides may be used in Southern or Northern analysis, polymerase
 CC chain reaction (PCR), or in enzyme-linked immunosorbent assays (ELISA).
 XX
 SQ Sequence 274 AA;
 Query Match 100.0%; Score 1386; DB 21; Length 274;
 Best Local Similarity 100.0%; Pred. No. 4.7e-131;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAAAEPSRRVGFVGAGRMAGTAAGLIRAGKVEAQHILASAPTRNLCHFQALGCRTH 60
 DB 1 MAAAEPSRRVGFVGAGRMAGTAAGLIRAGKVEAQHILASAPTRNLCHFQALGCRTH 60
 QY 61 SNOEVLQSCLLVIFATKPHVLPVLAEPVPTTEHILSVAGVSLTLEELLPPNTRV 120
 DB 61 SNOEVLQSCLLVIFATKPHVLPVLAEPVPTTEHILSVAGVSLTLEELLPPNTRV 120
 QY 121 LRVLPNLPCVVOEGAIWARGRHVGSSETKLQHLLEACGRCEVPPEAYVDIHTGLSGG 180
 DB 121 LRVLPNLPCVVOEGAIWARGRHVGSSETKLQHLLEACGRCEVPPEAYVDIHTGLSGG 180
 QY 181 VAFVCAFSEALAEAGVAKMGMPSSLAHRTAAQTLLGTAKMLLHEGQHPAQLRSDVCPGGT 240
 DB 181 VAFVCAFSEALAEAGVAKMGMPSSLAHRTAAQTLLGTAKMLLHEGQHPAQLRSDVCPGGT 240
 QY 241 TIYGLHALLEOQGLRAATMSAVEATCRAKELSRK 274
 DB 241 TIYGLHALLEOQGLRAATMSAVEATCRAKELSRK 274

RESULT 2
 AAB95591
 ID AAB95591 standard; Protein; 274 AA.
 XX
 AC AAB95591;
 XX
 XX 26-JUN-2001 (first entry)
 XX
 XX Human protein sequence SEQ ID NO:18269.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 XX Homo sapiens.
 XX
 XX EP1074617-A2.
 XX
 XX 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-0116126.
 XX
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 XX full-length cDNAs -
 PS Claim 8; SEQ ID 18269; 2537pp + CD ROM; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 274 AA;
 Query Match 100.0%; Score 1386; DB 22; Length 274;
 Best Local Similarity 100.0%; Pred. No. 4.7e-131;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAAAEPSRRVGFVGAGRMAGTAAGLIRAGKVEAQHILASAPTRNLCHFQALGCRTH 60
 DB 1 MAAAEPSRRVGFVGAGRMAGTAAGLIRAGKVEAQHILASAPTRNLCHFQALGCRTH 60

QY 61 SNOEVLOSCLLVIFATKPHVLPVLAELVAPVVTTEHILSVAAAGVSLSTLEELLPPNTRV 120
|||||
DB 61 SNOEVLOSCLLVIFATKPHVLPVLAELVAPVVTTEHILSVAAAGVSLSTLEELLPPNTRV 120
QY 121 LRVLPNLPVVOEGAIVMARGHRVGSSETKLLQHLLEACGRCCEVPEAYVDIHTGLSGSG 180
|||||
DB 121 LRVLPNLPVVOEGAIVMARGHRVGSSETKLLQHLLEACGRCCEVPEAYVDIHTGLSGSG 180
QY 181 VAFVCAFSEALAEAGAVKMGMPSSLAHRIAAGTLLGTAKMLLHGGHQAOLRSDVCTPGGT 240
|||||
DB 181 VAFVCAFSEALAEAGAVKMGMPSSLAHRIAAGTLLGTAKMLLHGGHQAOLRSDVCTPGGT 240
QY 241 TIYGLHALEQGLRAATMSAVEAATCRAKELSRK 274
|||||
DB 241 TIYGLHALEQGLRAATMSAVEAATCRAKELSRK 274

RESULT 3
AAG79743
ID AAG79743 standard; Protein: 274 AA.
XX
AC AAG79743;
XX
DT 18-MAR-2003 (first entry)
XX
DE P5CR related polypeptide #2.
XX
KW Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
KW Pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.
XX
OS Homo sapiens.
XX
PN WO200299043-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US17319.
XX
PR 05-JUN-2001; 2001US-296080P.
PR 10-OCT-2001; 2001US-328509P.
XX
PA (EXEL-) EXELIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Engst S;
XX
DR WPI; 2003-140606/13.
XX
PS Claim 13; Page 48-49; 53pp; English.
XX
CC The sequences given in AAG79742-46 represent pyrroline 5 carboxylate
CC reductase (P5CR) related polypeptide sequences. P5CR catalyses the
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
CC proline. These sequences may be used in the method of the invention
CC for identifying a candidate p53 pathway modulating agent. The method
CC comprises providing an assay system comprising a purified P5CR
CC polypeptide or nucleic acid, or a functionally active fragment or
CC derivative, contacting the assay system with a test agent, where the
CC system provides a reference activity, and detecting a test agent-biased
CC activity of the assay system. The method of the present invention is
CC useful for the diagnosis and treatment of disorders associated with
CC defects in the p53 pathway, such as cancer of the breast, colon,
CC kidneys, lung and ovary.
XX
SQ Sequence 274 AA;

Query Match 100.0%; Score 1386; DB 24; Length 274;
Best Local Similarity 100.0%; Pred. No. 4,7e-131;

Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAABESPVRVGFVAGRMAGAIAGLIRAGKVEAQHILLASAPTDRNLCHFQALGCRTH 60
|||||
DB 1 MAAABESPVRVGFVAGRMAGAIAGLIRAGKVEAQHILLASAPTDRNLCHFQALGCRTH 60
QY 61 SNOEVLOSCLLVIFATKPHVLPVLAELVAPVVTTEHILSVAAAGVSLSTLEELLPPNTRV 120
|||||
DB 61 SNOEVLOSCLLVIFATKPHVLPVLAELVAPVVTTEHILSVAAAGVSLSTLEELLPPNTRV 120
QY 121 LRVLPNLPVVOEGAIVMARGHRVGSSETKLLQHLLEACGRCCEVPEAYVDIHTGLSGSG 180
|||||
DB 121 LRVLPNLPVVOEGAIVMARGHRVGSSETKLLQHLLEACGRCCEVPEAYVDIHTGLSGSG 180
QY 181 VAFVCAFSEALAEAGAVKMGMPSSLAHRIAAGTLLGTAKMLLHGGHQAOLRSDVCTPGGT 240
|||||
DB 181 VAFVCAFSEALAEAGAVKMGMPSSLAHRIAAGTLLGTAKMLLHGGHQAOLRSDVCTPGGT 240
QY 241 TIYGLHALEQGLRAATMSAVEAATCRAKELSRK 274
|||||
DB 241 TIYGLHALEQGLRAATMSAVEAATCRAKELSRK 274

RESULT 4
AAG79745
ID AAG79745 standard; Protein: 274 AA.
XX
AC AAG79745;
XX
DT 18-MAR-2003 (first entry)
XX
DE P5CR related polypeptide #4.
XX
KW Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
KW Pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.
XX
OS Homo sapiens.
XX
PN WO200299043-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US17319.
XX
PR 05-JUN-2001; 2001US-296080P.
PR 10-OCT-2001; 2001US-328509P.
XX
PA (EXEL-) EXELIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Engst S;
XX
DR WPI; 2003-140606/13.
XX
PS Claim 13; Page 50-52; 53pp; English.
XX
CC The sequences given in AAG79742-46 represent pyrroline 5 carboxylate
CC reductase (P5CR) related polypeptide sequences. P5CR catalyses the
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
CC proline. These sequences may be used in the method of the invention
CC for identifying a candidate p53 pathway modulating agent. The method
CC comprises providing an assay system comprising a purified P5CR
CC polypeptide or nucleic acid, or a functionally active fragment or
CC derivative, contacting the assay system with a test agent, where the
CC system provides a reference activity, and detecting a test agent-biased
CC activity of the assay system. The method of the present invention is
CC useful for the diagnosis and treatment of disorders associated with
CC defects in the p53 pathway, such as cancer of the breast, colon,

The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 343 AA;

Query Match 42.6%; Score 590.5; DB 23; Length 343;
Best Local Similarity 45.1%; Pred. No. 8,9e-51;
Matches 125; Conservative 49; Mismatches 96; Indels 7; Gaps 3;

QY 1 MAAPSPRRVGVGAGRMAGATAGLIRAGKVEAQHILASAPTRNLCHFQAL---GCR 57
DB 17 LGQSPDSMSYGFAGAGLAFALAKGFTAAAGVLAHAKIMASSP-DMDLATVSAIRKMGVK 75
QY 58 TTHSNOEVLQSLAVIFATKPHVLPVLAELVAPVVTTEHILVSAAGVSLTLELLP-- 115
DB 76 LTPINKETVQMSDVLFLAVRPHIIPFLDEIGADIEDRHIVVSCAAGVTISSIEKKLSAF 135
QY 116 -PNTRLVRLNPLPCVVEGAIVMARGHRVGSSETKLQHLLEACGRCCEVPEAYVDIHT 174
DB 136 RPAPRVIRCMNTPPVVREGATVYATGTHAQVEDGRLEQMLSSVSGFCTEVEDLIDAVT 195
QY 175 GLSGSGVAFVCAFSEALAEAGVAKMGMPSSLAHRIAOTLLGTAKMLLHEGQHPAQLRSDV 234
DB 196 GLSGSGPAYAFTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPQLKDNV 255
QY 235 CTGPGTTIYGLHAEOGGLRAATMSAVEAATCAKREL 271
DB 256 SSPGATIHVLHVESGGRSLLINAVEASCIRTREL 292

RESULT 7

AAB74780
ID AAB74780 standard; Protein; 319 AA.

XX AAB74780;

XX 06-JUN-2001 (first entry)

XX Human P5CR protein.

XX Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.

XX Homo sapiens.

XX CN1274728-A.

PD 29-NOV-2000.
XX 25-MAY-1999; 99CN-0107071.
XX 25-MAY-1999; 99CN-0107071.
XX (UYFU-) UNIV FUDAN.
XX Yu L, Fu Q, Zhang H;
XX WPI; 2001-2111749/22.

XX New human protein and its code sequence, preparation and application -
XX Example 2; Fig 1; 20pp; Chinese.

XX The present invention describes a human protein designated Py-CR, which
XX is homologous to the human pyrroline-5'-carboxylate reductase (P5CR)
XX (EC1.5.1.2). The present invention also describes methods for the
XX application and production process of the Py-CR polynucleotide and
XX protein sequences. The present sequence represents the human P5CR
XX protein which is used in comparison with the human Py-CR protein in
XX an example from the present invention.

XX SQ Sequence 319 AA;

Query Match 42.2%; Score 584.5; DB 22; Length 319;
Best Local Similarity 46.4%; Pred. No. 3.2e-50;
Matches 124; Conservative 47; Mismatches 89; Indels 7; Gaps 3;

QY 11 VGFVGAAGMAGATAGLIRAGKVEAQHILASAPTRNLCHFQAL---GCRTHSNOEVLQ 67
DB 3 VGFIGAGLAFALAKGFTAAAGVLAHAKIMASSP-DMDLATVSAIRKMGVKLTPHKNETVQ 61
QY 68 SCLLVIFATKPHVLPVLAELVAPVVTTEHILVSAAGVSLTLELLP---PNTRLVRL 124
DB 62 HSDVFLAVKPHIIPFLDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRAPVIRCM 121
QY 125 PNLPCVVEGAIVMARGHRVGSSETKLQHLLEACGRCCEVPEAYVDIHTGLSGSVAFV 184
DB 122 TNPVVVREGATVYATGTHAQVEDGRLEQMLSSVSGFCTEVEDLIDAVTGLSGSPAYA 181
QY 185 CAFSEALAEAGVAKMGMPSSLAHRIAOTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
DB 182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPQLKDNVSSPGGATTHA 241
QY 245 LHAEQGLRAATMSAVEAATCAKREL 271
DB 242 LHVLESGGRSLLINAVEASCIRTREL 268

RESULT 8

AAG79742
ID AAG79742 standard; Protein; 319 AA.

XX AAG79742;

XX 18-MAR-2003 (first entry)

XX P5CR related polypeptide #1.

XX Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
XX pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 1..253

XX /label= P5CR domain

XX WO200299043-A2.

XX 12-DEC-2002.

```

XX PF 03-JUN-2002; 2002WO-US17319.
XX OS
XX PN WO200299043-A2.
XX PR 10-OCT-2001; 2001US-296080P.
XX PD 12-DEC-2002.
XX PA (EXEL-) EXELIS INC.
XX PF 03-JUN-2002; 2002WO-US17319.
XX PR 05-JUN-2001; 2001US-296080P.
XX PR 10-OCT-2001; 2001US-328509P.
XX PA (EXEL-) EXELIS INC.
XX PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX PI Engst S;
XX DR WPI; 2003-140606/13.
XX PT
XX PT Identifying p53 pathway modulating agents with p53 genes, useful for
XX PT the diagnosis and treatment of disorders associated with defects in the
XX PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
XX PT ovary.
XX PS Claim 13; Page 46-48; 53pp; English.
XX CC The sequences given in AAG79742-46 represent pyrroline 5 carboxylate
XX CC reductase (P5CR) related polypeptide sequences. P5CR catalyses the
XX CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
XX CC proline. These sequences may be used in the method of the invention
XX CC for identifying a candidate p53 pathway modulating agent. The method
XX CC comprises providing an assay system comprising a purified p5CR
XX CC polypeptide or nucleic acid, or a functionally active fragment or
XX CC derivative, contacting the assay system with a test agent, where the
XX CC system provides a reference activity, and detecting a test agent-biased
XX CC activity of the assay system. The method of the present invention is
XX CC useful for the diagnosis and treatment of disorders associated with
XX CC defects in the p53 pathway, such as cancer of the breast, colon,
XX CC kidneys, lung and ovary.
XX CC Sequence 319 AA;
XX QY 11 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTDRLNLCFQAL---GCRTHSNQEVLO 67
XX DB 11 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTDRLNLCFQAL---GCRTHSNQEVLO 67
XX QY 3 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTDRLNLCFQAL---GCRTHSNQEVLO 67
XX DB 3 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTDRLNLCFQAL---GCRTHSNQEVLO 67
XX QY 68 SCLLVIFATKPHVLPVLAEPVVTTEHLVSVAGVSLSTLELLP---PNTRVLRVL 124
XX DB 68 SCLLVIFATKPHVLPVLAEPVVTTEHLVSVAGVSLSTLELLP---PNTRVLRVL 124
XX QY 62 HSDVLFVAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAFVIRCM 121
XX DB 62 HSDVLFVAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAFVIRCM 121
XX QY 125 PNLPCVQVQEGATVMARGHVGSSSETKLLQHLLEACGRCVEEYVAYVDIHTGLSGGVAFV 184
XX DB 125 PNLPCVQVQEGATVMARGHVGSSSETKLLQHLLEACGRCVEEYVAYVDIHTGLSGGVAFV 184
XX QY 122 TNPVVRREGATVYATGTHAQVEDGRIMEQLLSTVGFCFTEVEDLIDAVTGLSGGPAYA 181
XX DB 122 TNPVVRREGATVYATGTHAQVEDGRIMEQLLSTVGFCFTEVEDLIDAVTGLSGGPAYA 181
XX QY 185 CAFSEALAEAGVAKMGMPSSLAHRIAQAOTLLGTAKMLLHGGHQAQLRSDVCTPGGTTIYG 244
XX DB 185 CAFSEALAEAGVAKMGMPSSLAHRIAQAOTLLGTAKMLLHGGHQAQLRSDVCTPGGTTIYG 244
XX QY 182 FTALDALADGGVKMGLPRFLAVRLCAQALLGNAKMLLHSEHQPGLKDNVSSPGGATIHA 241
XX DB 182 FTALDALADGGVKMGLPRFLAVRLCAQALLGNAKMLLHSEHQPGLKDNVSSPGGATIHA 241
XX QY 245 LHLEOGGLRAATMSAVEAATCRAKEL 271
XX DB 245 LHLEOGGLRAATMSAVEAATCRAKEL 271
XX QY 242 LHVLESGGFRSLLINAVEASCIRTREL 268
XX DB 242 LHVLESGGFRSLLINAVEASCIRTREL 268

RESULT 9
AAG79744
ID AAG79744 standard; Protein; 319 AA.
XX AC AAG79744;
XX DT 18-MAR-2003 (first entry)
XX DE P5CR related polypeptide #3.
XX KW Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
XX KW pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.

```

DT 15-JAN-2002 (first entry)
XX Human delta 1-pyrroline-5-carboxylate reductase homologue (P5CRH).
DE
XX Human; delta 1-pyrroline-5-carboxylate reductase; P5CR; cystic fibrosis;
KW osteoporosis; neuronal disorder; gene therapy; akathisia; drug screening;
KW actinic keratosis; Alzheimer's disease; amyotrophic lateral sclerosis;
KW connective tissue disorder; myocardial fibrosis; cell proliferation;
KW arteriosclerosis bursitis; cancer; amnesia; neuroprotective; cytostatic;
KW antiarteriosclerotic; osteopathic; cardiac.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..18
FT /label= Signal_peptide
FT Protein 19..320
FT /label= Mature_human P5CRH
FT Misc-difference 14
FT /note= "This amino acid residue is absent in the
FT sequence shown in sequence listing of the specification"
FT Misc-difference 15
FT /note= "This amino acid residue is absent in the
FT sequence shown in sequence listing of the specification"
FT Misc-difference 16
FT /note= "This amino acid residue is absent in the
FT sequence shown in sequence listing of the specification"
FT Misc-difference 18
FT /note= "This amino acid residue is absent in the
FT sequence shown in sequence listing of the specification"
FT Misc-difference 188
FT /note= "This amino acid residue is absent in the
FT sequence shown in sequence listing of the specification"
FT Misc-difference 189
FT /note= "This amino acid residue is absent in the
FT sequence shown in sequence listing of the specification"
XX
XX US6268192-B1.
XX
XX 31-JUL-2001.
XX
XX 05-MAY-2000; 2000US-0565910.
XX
XX 18-JUN-1998; 98US-0099676.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Hillman JL, Corley NC, Baughn MR;
XX WPI; 2001-647056/74.
XX N-PSDB; AAD20894.
XX
XX New delta 1-pyrroline-5-carboxylate reductase polypeptides and
XX polynucleotides, useful for diagnosing, treating and preventing
XX neuronal disorders, connective tissue disorders or disorders of cell
XX proliferation
XX
XX Claim 1; Fig 2; 32pp; English.
XX
XX The present invention relates to delta 1-pyrroline-5-carboxylate
XX reductase (P5CR) polypeptides and polynucleotides. The P5CR gene is used
XX in gene therapy. The P5CR nucleic acid and amino acid sequences are
XX useful in the diagnosis, treatment and prevention of neuronal disorders
XX (e.g. akathisia, Alzheimer's disease, amnesia, or amyotrophic lateral
XX sclerosis), connective tissue disorders (e.g. cystic fibrosis, myocardial
XX fibrosis, or osteoporosis) and disorders of cell proliferation (e.g.
XX actinic keratosis, arteriosclerosis bursitis, cancers). P5CR may also be
XX used for screening libraries of compounds in various drug screening
XX techniques. The present sequence is human delta 1-pyrroline-5-carboxylate
XX reductase homologue (P5CRH).
XX
XX Sequence 320 AA;

Query Match 41.7%; Score 577.5; DB 22; Length 320;
Best Local Similarity 46.4%; Pred. No. 1.6e-49;
Matches 124; Conservative 48; Mismatches 88; Indels 7; Gaps 3;
QY 11 VGFVAGRMAGAIAGGLIRAGKVEAQHILASAPTRDNLCHFOAL---GCRTHSNQEVQLQ 67
DB 3 VGFICAGQLAYALARGFTAGILSAHKIIASSP-EMNLPTVSALRKMGNLRSNKETVK 61
QY 68 SCLLVIFATKPHVLPVLAELVAVPVVTTTHILVSVAGVSLTLRELL---PNTRVLRLVL 124
DB 62 HSDVLFVAVKPHIIPFIDEIGADVQARHIVVSCAAQTIISSVEKKLMAFQAPKVIKCM 121
QY 125 PNLPCVQOEGATVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGGVAFV 184
DB 122 TMTVVVOEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGGPAY 181
QY 185 CAFSEALAEAGVKGMPSSLAHRIAQTLLGTAKMLLHGGHPAQLRSDVCTPGTTIYG 244
DB 182 FMDALADGGVKGMLPRRLAIQLGAQALLGAAKMLLDSEQHPCOLKDNVCSPPGATIIHA 241
QY 245 LHALEGGGLRAATMSAVEAATCRAKEL 271
DB 242 LHFLESGGFRSLINAVEASCIRTREL 268
RESULT 11
ABB61855
ID ABB61855 standard; Protein; 273 AA.
XX
XX ABB61855;
XX
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 12357.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL05958.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions
XX
XX Disclosure; SEQ ID NO 12357; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX

CC	sclerosis), connective tissue disorders (e.g. cystic fibrosis, myocardial
CC	fibrosis, or osteoporosis) and disorders of cell proliferation (e.g.
CC	actinic keratosis, arteriosclerosis bursitis, cancers). P5CR may also be
CC	used for screening libraries of compounds in various drug screening
CC	techniques. The present sequence is human delta 1-pyrroline-5-carboxylate
CC	reductase (P5CR).
XX	CC
SQ	Sequence 315 AA;
	Query Match 40.4%; Score 560.5; DB 22; Length 315;
	Best Local Similarity 46.1%; Pred.No.8.3e+48;
	Matches 123; Conservative 46; Mismatches 87; Indels 11; Gaps 5;
QY	11 VGFVCGAGMAGAQAOLIRAGKVEAOHILASAPTRNLCHFQAL--GCRTHSNOEVLQ 67
DB	: :: :: :: : : :: : : : : :
	3 VGFIAGCOL--AFAKGTAGVALAAHKIMASSP-DMDLATVSALLRMGWKLTPHNKETVQ 59
QY	68 SCLLVFATKPHEVLPAVLAIEVAPVVTEHLINLSVAAGVSLSTLEELLP---PNTRLRLVL 124
DB	:: :: :: : : : : : : : :
DB	60 HSDVFLAVKKPHIITPFIDIGADIEDRHVTWSCAAGVTISSEIKLSAFRAPVRICM 119
QY	125 PNLCPCVGQEGAIVMARGHRHVGSSETKLQLLBEACGRCEEPVEAYVDHTGLSGSGVAFV 184
DB	:: :: :: : : : : : : : :
DB	120 INTPTPVREGATVTATGTHAQVEDGRLMEOCLLTGVCTEVEDLDIVATGLSGGPAY - 178
QY	185 CAFSEALAEGAVKNMGMPSSSLAHRIAQTLLGTAKMLLHEGOHPAQLRSDVCITPGCTTIYG 244
DB	:: :: :: : : : : : : : :
DB	179 -AFTALDADGGVKMKPRRLVARLVGAQALIGAAMKLLHSEQHPQLKONVSPGGATHA 237
QY	245 LHALFQQGLRAATNSAVEAATCAKREL 271
DB	:: :: :: : : : : : : : :
DB	238 LHVLSEGGRFSLLINAWEASCIRTRE 264

RESULT 13	
AAB74779	
ID	AAB74779 standard; Protein; 320 AA.
XX	
AC	AAB74779;
XX	
DT	06-JUN-2001 (first entry)
XX	
DE	Human Py-CR protein SEQ ID NO:4.
XX	
KW	Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.
XX	
OS	Homo sapiens.
XX	
PN	CN1274728-A.
XX	
PD	29-NOV-2000.
XX	
PF	25-MAY-1999; 99CN-0107071.
XX	
PR	25-MAY-1999; 99CN-0107071.
XX	
PA	{UYFU-} UNIV FUDAN.
XX	
PI	Yu L, Fu Q, Zhang H;
XX	
DR	WPI: 2001-211749/22.
DR	N-PSDB; AAF81847.
XX	
XX	New human protein and its code sequence, preparation and application -
PT	
PS	Claim 4; Page 16; 20pp; Chinese.

The present invention describes a human protein designated Py-CR, which is homologous to the human pyrroline-5'-carboxylate reductase (P5CR) (EC1.5.1.2). The present invention also describes methods for the application and production process of the Py-CR polynucleotide and protein sequences. The present sequence represents the human Py-CR protein as given in the present invention.

```
XX SQ Sequence 320 AA;
Query Match 40.4%; Score 560.5; DB 22; Length 320;
Best Local Similarity 45.7%; Pred. No. 8.5e-48;
Matches 122; Conservative 48; Mismatches 90; Indels 7; Gaps 3;

QY 11 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTRDNLCHFQAL---GCRTHSNQEVLIQ 67
DB 3 VGFIGAGQLANALARGFTAGILSAHKIIASSP-EMNLPTVSLRKMGNVLRNSKETVK 61
QY 68 SCLLVIFATKPHVLPVLAEPVVTTEHILVSAAGVSLSTLEELL---PPNTRVLRVL 124
DB 62 HSDVFLAVKHHIIPFLDEIGADVQARHIVVSCAAGVTISSVEKKLMFAQPAKPVIRCM 121
QY 125 PNLPCVVOEGAIVMARGHRVGSSETKLLQHLLEACRCREVPAYVDIHTGLSGSVAFV 184
DB 122 TNPVVVVOEGATVYATGTHALVEDGQLLEQMSVGVCFTEVEDLIDAVTGLSGSGPAY 181
QY 185 CAFSEALAEAGVAKMGMPSSLAHRIAATLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
DB 182 FMALDALADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSQHPCLKDNVCSPPGATIIHA 241
QY 245 LHALEQGGGLRAATMSAVEAATCRAKEL 271
DB 242 LHFLESGGFRSLLINAVEASCIRTREL 268

RESULT 14
AAB20584
ID AAB20584 standard; Protein: 314 AA.
XX AC AAB20584;
XX DT 13-DEC-2000 (first entry)
XX DE Human delta 1-pyrroline-5-carboxylate reductase homologue protein.
XX KW Human; delta 1-pyrroline-5-carboxylate reductase homologue; P5CRH;
XX KW neotropic; neuroprotective; osteopathic; cytotostatic; antiparkinsonian;
XX KW cardiant; osteopathic; ophthalmological; hepatotropic; gene therapy;
XX KW neuronal disorder; connective tissue disorder; cell proliferation;
XX KW akathisia; Alzheimer's disease; amnesia; dementia; Parkinson's disease;
XX KW cystic fibrosis; myocardial fibrosis; osteoporosis; Marfan syndrome;
XX KW cirrhosis, leukaemia, lymphoma, sarcoma, bone marrow cancer.
XX OS Synthetic.
XX PN US6100075-A.
XX PD 08-AUG-2000.
XX PF 18-JUN-1998; 98US-0099676.
XX PR 18-JUN-1998; 98US-0099676.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Hillman JL, Baughn MR, Corley NC;
XX DR WPI; 2000-548292/50.
XX DR N-PSDB; AAA88073.
XX PT New human delta 1-pyrroline-5-carboxylate reductase homolog (P5CRH) and
XX PT polynucleotides encoding P5CRH, useful for diagnosing, treating or
XX PT preventing neuronal or connective tissue disorders, or disorders of
XX PS cell proliferation
XX PS Claim 1; Fig 1A-E; 32pp; English.
XX CC The present sequence represents human delta 1-pyrroline-5-carboxylate
XX CC reductase homologue (P5CRH). P5CRH has neotropic, neuroprotective,
XX CC osteopathic, cytotostatic, antiparkinsonian, cardiant, osteopathic,
```

```
CC ophthalmological and hepatotropic activities, and can be used in gene
CC therapy. P5CRH and the polynucleotides encoding it are useful for
CC diagnosing, treating or preventing neuronal disorders, connective tissue
CC disorders, or disorders of cell proliferation. These may be applied to
CC any subject, e.g. dogs, cats, cows, horses, rabbits, monkeys, or
CC preferably humans. The polynucleotide sequences may be used to detect
CC and quantitate gene expression in biopsied tissues where expression of
CC the polypeptide may correlate with the disease. The diagnostic assay may
CC be used to determine absence, presence, and excess expression of the
CC polypeptide, and to monitor regulation of the polypeptide levels during
CC therapeutic intervention. The polypeptide or its fragments are useful
CC for treating or preventing neuronal disorder (e.g. akathisia,
CC Alzheimer's disease, amnesia, dementia or Parkinson's disease),
CC connective tissue disorder (e.g. cystic fibrosis, myocardial fibrosis,
CC osteoporosis or Marfan syndrome), or a disorder of cell proliferation
CC (e.g. cirrhosis, leukaemia, lymphoma, sarcoma or bone marrow cancer).
CC The polypeptide or its fragments are also useful for screening
CC libraries of compounds in any of several drug-screening techniques.
XX SQ Sequence 314 AA;
Query Match 38.8%; Score 537.5; DB 21; Length 314;
Best Local Similarity 45.3%; Pred. No. 1.7e-45;
Matches 121; Conservative 45; Mismatches 88; Indels 13; Gaps 5;

QY 11 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTRDNLCHFQAL---GCRTHSNQEVLIQ 67
DB 3 VGFIGAGQ----LAYRFTAAGILSAHKIIASSP-EMNLPTVSLRKMGNVLRNSKETVK 57
QY 68 SCLLVIFATKPHVLPVLAEPVVTTEHILVSAAGVSLSTLEELL---PPNTRVLRVL 124
DB 58 HSDVFLAVKHHIIPFLDEIGADVQARHIVVSCAAGVTISSVEKKLMFAQPAKPVIRCM 117
QY 125 PNLPCVVOEGAIVMARGHRVGSSETKLLQHLLEACRCREVPAYVDIHTGLSGSVAFV 184
DB 118 TNPVVVVOEGATVYATGTHALVEDGQLLEQMSVGVCFTEVEDLIDAVTGLSGSGPAY- 176
QY 185 CAFSEALAEAGVAKMGMPSSLAHRIAATLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
DB 177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSQHPCLKDNVCSPPGATIIHA 235
QY 245 LHALEQGGGLRAATMSAVEAATCRAKEL 271
DB 236 LHFLESGGFRSLLINAVEASCIRTREL 262

RESULT 15
AAU99322
ID AAU99322 standard; Protein: 314 AA.
XX AC AAU99322;
XX DT 24-SEP-2002 (first entry)
XX DE Human delta 1-pyrroline-5-carboxylate reductase homologue protein.
XX KW Human; neotropic; neuroprotective; cytotostatic; enzyme; P5CRH;
XX KW delta 1-pyrroline-5-carboxylate reductase homologue; proline;
XX KW glutamate biosynthesis; collagen; bone; interstitial fluid;
XX KW neuronal disorder; Alzheimer's disease; Parkinson's disease;
XX KW connective tissue disorder; cystic fibrosis; osteoporosis;
XX KW cell proliferation; atherosclerosis; cancer.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 13..15
XX FT /note= "Encoded by TATGCTCTGGCGGGGCTTC"
XX FT Misc-difference 183..184
XX FT /note= "Encoded by GCATTGGCTGAT"
XX FT Region 212..242
XX FT /note= "Delta 1-pyrroline-5-carboxylate reductase
XX FT signature motif"
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 30 Seconds
(without alignments)
386.439 Million cell updates/sec

Title: US-09-806-536A-14

Perfect score: 1386

Sequence: 1 MAAAEPSRRVGVGAGRMA.....AATMSAVEATCRAKELSRK 274

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	560.5	40.4	315	3	US-09-099-676-3
2	560.5	40.4	315	3	US-09-565-910-3
3	537.5	38.8	314	3	US-09-099-676-1
4	537.5	38.8	314	3	US-09-565-910-1
5	345.5	24.9	472	4	US-09-252-991A-17011
6	327.5	23.6	289	4	US-09-107-532A-7169
7	321	23.2	281	4	US-09-328-352-8170
8	297.5	21.5	304	1	US-08-665-716-2
9	269	19.4	282	4	US-09-134-001C-3155
10	179.5	13.0	271	4	US-09-328-352-6802
11	134	9.7	171	4	US-09-252-991A-20711
12	116	8.4	144	4	US-08-936-165A-445
13	108.5	7.8	409	4	US-09-252-991A-20236
14	97.5	7.0	1238	4	US-09-252-991A-26363
15	97	7.0	398	4	US-09-252-991A-20748
16	96.5	7.0	303	4	US-09-328-352-4879
17	96.5	7.0	516	4	US-09-252-991A-17933
18	96.5	7.0	811	4	US-09-252-991A-22006
19	91.5	6.6	436	3	US-08-669-378-4
20	91	6.6	280	4	US-09-328-352-4975
21	91	6.6	6396	4	US-09-410-551B-72
22	90.5	6.5	851	4	US-09-252-991A-19645
23	88	6.3	342	4	US-09-364-230-6
24	87.5	6.3	221	4	US-09-252-991A-23434
25	87.5	6.3	537	4	US-09-252-991A-26178
26	87.5	6.3	565	3	US-09-008-481A-6
27	87.5	6.3	565	3	US-09-195-666A-5

28 87.5 6.3 565 3 US-09-195-666A-49 Sequence 49, Appl
29 87.5 6.3 565 3 US-09-309-592-6 Sequence 6, Appl
30 87.5 6.3 565 4 US-09-635-705-5 Sequence 5, Appl
31 87.5 6.3 565 4 US-09-635-705-49 Sequence 49, Appl
32 87.5 6.3 565 4 US-09-634-858A-5 Sequence 5, Appl
33 87.5 6.3 565 4 US-09-634-858A-49 Sequence 49, Appl
34 87.5 6.3 565 4 US-08-869-927C-5 Sequence 5, Appl
35 87.5 6.3 565 4 US-08-869-927C-49 Sequence 49, Appl
36 87.5 6.3 762 4 US-09-252-991A-28078 Sequence 28078, A
37 87.5 6.3 1024 4 US-09-562-737-44 Sequence 44, Appl
38 87 6.3 528 3 US-08-928-213B-8 Sequence 8, Appl
39 87 6.3 949 3 US-09-196-387-10 Sequence 10, Appl
40 87 6.3 949 4 US-09-841-835-10 Sequence 10, Appl
41 87 6.3 1327 3 US-09-196-387-2 Sequence 2, Appl
42 87 6.3 1327 4 US-09-841-835-2 Sequence 2, Appl
43 86.5 6.2 418 4 US-09-328-352-5304 Sequence 5304, Ap
44 86.5 6.2 510 4 US-09-252-991A-25949 Sequence 25949, A
45 86.5 6.2 578 4 US-09-252-991A-25113 Sequence 25113, A

ALIGNMENTS

RESULT 1

US-09-099-676-3
; Sequence 3, Application US/09099676
; Patent No. 6100075
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
; TITLE OF INVENTION: HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,676
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerione, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0532 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-855-0572
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 189498

Query Match 40.4%; Score 560.5; DB 3; Length 315;

[illegible]

```

RESULT 2
US-09-565-910-3
; Sequence 3, Application US/09565910
; Patent No. 6268192
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
; TITLE OF INVENTION: HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/565,910
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/099,676
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0532 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-855-0572
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 189498
; US-09-565-910-3

```

Query Match	40.4%;	Score 560.5;	DB 3;	Length 315;
Best Local Similarity	46.1%;	Pred. No. 1.6e-53;		
Matches 123;	Conservative 46;	Mismatches 87;	Indels 11;	Gaps 5
Qy	11	VGFVAGRMAGIAOGLIRAGKVEAOHILASAPATDRNLCHFOAL--GCRTHHSNOEVLQ 67		
Db	3	VGFIGAGQL--AFAGFTTAAGVLAHAKHIMASSP--DMDLATVSALRKMGYKLTIPHNKTVQ 59		
Qy	68	SCLLVIFATKPHVLPAVIAEAVPVVYTHEILVSVAAGVSGTLEELP---PNTRVLRVL 124		
Db	60	HSDFVLAVKPHIIPFIIDETGADIEDRHIVVSCAAGVTTISSIEKKLSAFRPAIPRVIRCM 119		
Qy	125	PNLPCVQBGATVMARGRHVSGSETKLOHLLACAGCEEPYAVVDIHTGLSGSGVAFV 184		
Db	120	TNTPVVREGATVYATGTHAQVEDGRLEMQLLSTVGCTEVEDLDVATVGLSSGSPAY- 178		
Qy	185	CAFSEALAEAVKMGMPSSLAHRTAAQTLLGTAKMLLHEGQHPAQLRSOVCTPGGTTIVG 244		
Db	179	-AFTALDAGGVKMGSLPRRLRAVRVAGQALLGAAMKLLHSEQHGPQLKDNVSSPGGATIIA 237		
Qy	245	LHAEQGGRLRAATMSAVEAATCAKEL 271		
Db	238	LHWLESGGFRSLIINAVEASCIIRTEL 264		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 38.8%; Score 537.5; DB 3; Length 314;
Best Local Similarity 45.3%; Pred. No. 5.7e-51;
Matches 121; Conservative 45; Mismatches 88; Indels 13; Gaps 5;
QY 11 VGFVAGRMAGAIAGGLIRAGKVEAQHILASAPTDRLNCHFOAL---GCRTHHSNOEVLQ 67
Db 3 VGFVAGQO---LAVRFTAAGILSAHKIIASSP-EMNLPTVSALRKMGNLTRSNKETVK 57
QY 68 SCLLVIFATKPHVLPVLAELAVPVVTTTEHILVSVAAAGVSLSTLEELL---PPNTRVLRLV 124
Db 58 HSDVLELAVKPHIIPFILDIGADVOARHIVVSCAAGVTISSVEKKLMAFOPAPKVIKRCM 117
QY 125 PNLPCVVOEGATVMAARGHVGSSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
Db 118 TNPVVOEGATVYATGTHALVEDQGLLEQLMSSVGFCTEVEDLIDAVTGLSGSGPAY- 176
QY 185 CAFSEALAGAVKMGMPSSLAHRIAQTLLGTAKMLLHEGOHPAQLRSDVCTPGGTTIYG 244
Db 177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPPGATIIHA 235
QY 245 LHAEQGLRAATMSAVEAATCRAKEL 271
Db 236 LHFLSGGFRSLINAVEASCIRTREL 262

RESULT 4
US-09-565-910-1
; Sequence 1, Application US/09565910
; Patent No. 6268192
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
; TITLE OF INVENTION: HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; APPLICATION DATA:
; APPLICATION NUMBER: US/09/565,910
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/099,676
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0532 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-855-0572
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNON01
; CLONE: 2278458

US-09-565-910-1
Query Match 38.8%; Score 537.5; DB 3; Length 314;
Best Local Similarity 45.3%; Pred. No. 5.7e-51;
Matches 121; Conservative 45; Mismatches 88; Indels 13; Gaps 5;
QY 11 VGFVAGRMAGAIAGGLIRAGKVEAQHILASAPTDRLNCHFOAL---GCRTHHSNOEVLQ 67
Db 3 VGFVAGQO---LAVRFTAAGILSAHKIIASSP-EMNLPTVSALRKMGNLTRSNKETVK 57
QY 68 SCLLVIFATKPHVLPVLAELAVPVVTTTEHILVSVAAAGVSLSTLEELL---PPNTRVLRLV 124
Db 58 HSDVLELAVKPHIIPFILDIGADVOARHIVVSCAAGVTISSVEKKLMAFOPAPKVIKRCM 117
QY 125 PNLPCVVOEGATVMAARGHVGSSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
Db 118 TNPVVOEGATVYATGTHALVEDQGLLEQLMSSVGFCTEVEDLIDAVTGLSGSGPAY- 176
QY 185 CAFSEALAGAVKMGMPSSLAHRIAQTLLGTAKMLLHEGOHPAQLRSDVCTPGGTTIYG 244
Db 177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPPGATIIHA 235
QY 245 LHAEQGLRAATMSAVEAATCRAKEL 271
Db 236 LHFLSGGFRSLINAVEASCIRTREL 262

RESULT 5
US-09-252-991A-17011
; Sequence 17011, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17011
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17011
Query Match 24.9%; Score 345.5; DB 4; Length 472;
Best Local Similarity 32.2%; Pred. No. 1.7e-29;
Matches 87; Conservative 49; Mismatches 131; Indels 3; Gaps 3;
QY 7 SPREVGFVAGRMAGAIAGGLIRAGKVEAQHILASAPTDRLNCHFOA-LGCRTHHSNOEV 65
Db 201 STPIAFIFAGNMAASLIGGL-RAQGVPAQIRASDPGAQRAKIAEGFAIDVVSNAEA 259
QY 66 LOSCLLVIFATKPHVLPVLAELAVPVVTTTEHILVSVAAAGVSLSTLEELLPPNTRVLRLV 125
Db 260 VADADVVLVSVPQAMKAVCALAPALKPEQLIVSIAAGIPCASLEAWLQGPVPVRCMP 319
QY 126 NLPVVOEGATVMAARGHVGSSSETKLLQHLLEACGRCEEV-PEAYVDIHTGLSGSGVAFV 184
Db 320 NTPALLRQAGSLYANAOVSAOQEQQLLSAVGIALWLDDEAQIDAVTAVSGSGPAY 379
QY 185 CAFSEALAGAVKMGMPSSLAHRIAQTLLGTAKMLLHEGOHPAQLRSDVCTPGGTTIYG 244
Db 380 FLLMQAMTDAGEKLGSLRSTASRLTLQALGAQMAALSSEVEPAELRRRTVSPNGTEAA 439
QY 245 LHAEQGLRAATMSAVEAATCRAKELSRK 274
Db 440 IKSFQANGFEALVEQALNNAASORSALAEQ 469

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: KORSN, ELLIOTT
REGISTRATION NUMBER: 32,705
REFERENCE/DOCKET NUMBER: 19453PV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-5493
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 304 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-665-716-2

Query Match 21.5%; Score 297.5; DB 1; Length 304;
Best Local Similarity 29.3%; Pred. No. 1.7e-24;
Matches 88; Conservative 52; Mismatches 121; Indels 39; Gaps 9;
3 AAEPSPRVGVFGAGRMAGATAOGLIRA-GKVEAOHILASAPTDRNLCHF--QALGCRTH 59
5 AKEGSELTVLIGCGTGIALSLGILASLDLHAPN-SQSETDETTSKLPTRFIACVRS 63
60 HSNQEVLSCL-----LVIFATKPHVLPVLAELVAPVVTTE-H 96
64 PKGAEKIKKALSPYKTPVKIIQSDNVATCREADVLLGCKPYMAEGILGEGMVDALKGK 123
97 ILVSVAAAGVSLTEELL---PN-----TRVLRVLPNLPVVOEGAIVMARGR-HVG 145
124 LLISILAGVPAEQIYGYWGTPTNPENKEGLCQVVRAMPNTASGIRESMVTIATSPPLS 183
146 SSETKLLQHLLEACGRCEVEPEAYVDIHTGLSGGVAFVCAFSALAEAGVAKMGMPSSLA 205
184 ATSSLLTIWFKRIGDVQLPAATMDASTALCGSPAFFALLLEAIDGAVMGLPRAEA 243
206 HRIAAQTLLGAKMLLEHGHQHPAQLRSDVCTPGGTTIYGLHALBOGGLRAATMSAVBAAT 265
244 ORMAAQTMKGAAGLVL-SGHPALLKDKVTPPGCTTIGGLMVLBEGGVGTVARAVREAT 302

RESULT 9

US-09-134-001C-3155
; Sequence 3155, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3155
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3155

Query Match 19.4%; Score 269; DB 4; Length 282;
Best Local Similarity 26.9%; Pred. No. 2.1e-21;
Matches 71; Conservative 54; Mismatches 137; Indels 2; Gaps 2;
13 FVGAGRMAGATAOGLIRAGKVEAQHI-LASAPTDRNLCHF-QALGCRTHSNQEVLSCL 70
16 FYGAGNMAQAFTGIINNNLNDIYLTNKSNEQALKSFAEKLGVNYSYDDEALKDAD 75
71 LVIFATKPHVLPVLAELVAPVVTTEHILVSVAAAGVSLTEELLPPNTRVLRVLPNLPV 130

Db 76 YVFLGTPKPHDFENLANRIRREHITNDNRFISIMAGLSIDYIROQLNTNPNLARIMPNTNAQ 135
QY 131 VOEGAIVMARGRHVGSGSETKLLQHLLEACGRCEVEPEAYVDIHTGLSGGVAFVCAFS 190
Db 136 VGHVGTGISFSNPNFDPKSKNEVDELINAFGSVTEVEEHLHQVTAITGSGSAPLYHVFEQ 195
QY 191 LAEGAVKMGMPSSIAHRIAAQTLLGTAKMLLEHGHQHPAQLRSDVCTPGGTTIYGLHALBQ 250
Db 196 YVKAGTGLGLERNQVEISIRNLIIGTSKMIERSDLSMSQLRKNTSKGGTTQAGLDALSQ 255
QY 251 GGLRAATMSAVEAATCRAKELSRK 274
Db 256 YDIVSMFEDCLGAAVNRSMELSHK 279

RESULT 10

US-09-328-352-6802
; Sequence 6802, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6802
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6802

Query Match 13.0%; Score 179.5; DB 4; Length 271;
Best Local Similarity 23.1%; Pred. No. 1.5e-11;
Matches 63; Conservative 51; Mismatches 124; Indels 35; Gaps 5;
13 FVGAGRMAGATAOGLIRAG-----KVEAQHILASAPTDRNLCHFQALGCRTH 60
11 FIGSSNLALIGLVKGFQREKINLIEVFENQIL-----K 50

QY 61 SNOEVLQSCLLVIFATKPHVLPVLAELVAPVVTTEHILVSVAAAGVSLTEELLPPNTRV 120
Db 51 OKOHEVKKADIVVLLDPKDLKAILAPLKKWL-ADKTIIVMMAGVNIQILMS-ITGSKKI 108
QY 121 LRVLPNLPVVOEGAIVMARGRHVGSGSETKLLQHLLEACGRCEV-PEAYVDIHTGLSGS 179
Db 109 IRVISNPPVLTYTGTHTVLIGSDYLEPLDKVEIETIYSATGRTYWANSESQSDAILALSGS 168
QY 180 GVAFVCAFSALAEAGVAKMGMPSSLAHRIAATLLGTAKMLLEHGHQHPAQLRSDVCTPGG 239
Db 169 GPAYFFYILDSMVYKTCVSMGLDKQFALDLILQNASGAVEMVRKSNVQPSLGCQKVTLANG 228
QY 240 TTYTGLHALBQGLRAATMSAVEAATCRAKELS 272
Db 229 ITESALRMFELGNLSDDIRLALKAAVHRSKEIN 261

RESULT 11

US-09-252-991A-20711
; Sequence 20711, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20711
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-20711

Query Match
Best Local Similarity 9.7%; Score 134; DB 4; Length 171;
Matches 29; Conservative 17; Mismatches 40; Indels 0; Gaps 0;

QY 189 EALAGAVKMPGSLAHRIAQTLTGAKMLLHGQHPAQLRSDVCTPGGTTIYGLHAL 248
Db 83 QAMTDAGEKGLGSLRETASRLTLQALGAQAQMSLSEVEPAELRRRTVSPNGTTEAATKSF 142
QY 249 EOGGLRAATMSAVAEATCRAKELSRK 274
Db 143 QANGFEALVEQALNAASORSLEAEQ 168

RESULT 12
US-08-936-165A-445
; Sequence 445, Application US/08936165A
; Patent No. 6348592
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 445:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-936-165A-445
```

```

Query Match
Best Local Similarity 8.4%; Score 116; DB 4; Length 144;
Matches 30; Conservative 24; Mismatches 56; Indels 2; Gaps 2;

QY 16 AGRMAGAIAGLIRAGKVEAQHI-LASAPTDRLNLCHF-QALGCRTHHSNQEVLSCLLVI 73
Db 12 AGNMAQAIPTGIINSNLDANDIYLTNKSNEQALKAFAEKLGVNYSYDDATLLKADAYVF 71
QY 74 FATKPHVLPVLAELAVPVVTTTEHILSVAAAGVSLSTLEELLPPNTRVLRVLP 125
Db 72 LGTKPHDFDALATRIKPHITKDXCFMSIMAGITDYIXOOLECONPFXARIMP 123

RESULT 13
US-09-252-991A-20236
; Sequence 20236, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20236
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-20236

Query Match
Best Local Similarity 7.8%; Score 108.5; DB 4; Length 409;
Matches 74; Conservative 30; Mismatches 102; Indels 75; Gaps 17;

QY 4 AEPSPRRV--GFVGAGRMA-GAIAOGLIRAGKVEAQHILASAPTDRLNLCHFQA-LGCRTT 59
Db 90 ALPGPRATLEGHPGGSGVPWGEYATGVIR-----RYPAATADNRKRFLFAQAHVPCRKA 143
QY 60 HSNQEVLSCLLVIATKPHVLPVLAELAVPV-VTTEHILSVAAAGVSLSTLEELLPPNT 118
Db 144 TRGWPF---CMATNATGAIRCLPAVLLHRRHPAQVTDHHRVVVGGVRLGVVQAVLLPBG 200
QY 119 RVLRLV---PNLPCV-----VOEGAIVMA--RGRHVG--SSTKLLQHL 156
Db 201 LEDVVVQHPGLVGVDPVPLHRAAYLARQAQGVGVKAGLPGEHVGIASCARLTRHQV 260
QY 157 EACGRCEEVPEAYVDIHTGLSGSVAFVCAFEALAEAGVK-MGMPSSIAHRIAQAOTLLG 215
Db 261 GVAGAAD-----DVHRVFLGIGI-----EVADDDQAIRLLG-----ARRIARQP--- 298
QY 216 TAKMLLHGHQHPAQLRSDVCTPGGTTIYGLHALEQGLRAA 256
Db 299 ----VHQGLGGAR-----PGQVAV----ALAVAGVRVA 323

RESULT 14
US-09-252-991A-26363
; Sequence 26363, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 98 seconds
(without alignments)
721.494 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSRRVGVGAGRNA.....AATMSAVEATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1386	100.0	274	4	Q9H896
2	1381	99.6	274	4	Q96HX4
3	1235	89.1	244	4	Q9N3N9
4	1163	83.9	274	11	Q9DCC4
5	1160	83.7	274	11	Q8R0P9
6	921	66.5	248	11	Q9D0X2
7	613	44.2	279	5	Q21544
8	585	42.2	270	5	Q8M070
9	577.5	41.7	320	4	Q96C36
10	571.5	41.2	320	11	Q922Q4
11	560.5	40.4	273	5	Q9VEJ3
12	560.5	40.4	309	11	Q922W5
13	555.5	40.1	274	2	Q8GE04
14	524.5	37.8	270	17	Q8P1P4
15	505.5	36.5	280	5	Q9V3F8
16	503	36.3	274	10	Q9AYM3

17	493.5	35.6	280	5	O96643	O96643 drosophila
18	480	34.6	270	16	O97E64	O97E64 clostridium
19	477	34.4	319	4	Q9Y5J4	Q9Y5J4 homo sapien
20	469	33.8	266	16	Q8Y9X2	Q8Y9X2 listeria mo
21	458	33.0	266	16	Q92EP4	Q92EP4 listeria in
22	457	33.0	284	10	Q8GT01	Q8GT01 oryza sativ
23	455.5	32.9	270	16	Q8YZH2	Q8YZH2 anabaena sp
24	438.5	31.6	266	2	O87725	O87725 clostridium
25	434.5	31.3	266	16	Q8XH52	Q8XH52 clostridium
26	430	31.0	266	16	Q8KG91	Q8KG91 chlorobium
27	429	31.0	264	16	Q8RA68	Q8RA68 thermoaer
28	426.5	30.8	272	4	Q9HBQ4	Q9HBQ4 homo sapien
29	425.5	30.7	305	16	Q8FKB0	Q8FKB0 escherichia
30	423.5	30.6	269	16	Q8XEB1	Q8XEB1 escherichia
31	421.5	30.4	269	16	Q8ZRF0	Q8ZRF0 salmonella
32	420.5	30.3	284	16	Q9X8G1	Q9X8G1 streptomyce
33	419.5	30.3	269	16	Q8Z8Z2	Q8Z8Z2 salmonella
34	418.5	30.2	262	5	Q8IDC6	Q8IDC6 plasmodium
35	416.5	30.1	272	2	Q9K308	Q9K308 bacillus ce
36	405	29.2	261	17	Q8U084	Q8U084 pyrococcus
37	404	29.1	174	5	Q8IN96	Q8IN96 drosophila
38	395	28.5	265	16	Q97R93	Q97R93 streptococc
39	394	28.4	283	3	Q96WX7	Q96WX7 emerigella
40	385	27.8	265	16	Q8DQ59	Q8DQ59 streptococc
41	371.5	26.8	256	16	Q8DS61	Q8DS61 streptococc
42	369	26.6	158	5	Q8MUT1	Q8MUT1 tigrilopus c
43	367	26.5	254	5	Q9NGS4	Q9NGS4 leishmania
44	364.5	26.3	314	10	Q8S488	Q8S488 phytophthor
45	363.5	26.2	273	16	Q8DJK6	Q8DJK6 synechococc

ALIGNMENTS

RESULT 1

Q9H896 ID Q9H896 PRELIMINARY; PRT; 274 AA.
AC Q9H896;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ13852.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RL "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK023314; BAB14721.1;
DR InterPro: IPR000304; P5CR.
DR Pfam: PF01089; P5CR; 1.
DR TIGRFAMS: TIGR00112; proc; 1.
DR PROSITE: PS00521; P5CR; 1.
KW Hypothetical protein.
SQ SEQUENCE 274 AA; 28663 MW; 846FDEC603F3B548 CRC64;

Query Match 100.0%; Score 1386; DB 4; Length 274;
Best Local Similarity 100.0%; Pred. No. 3.9e-103;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAAEPSRRVGVGAGRNAAGLIRAGKVEAQHILASAPTDRLNLCFQALGRTH 60
|||||
Db 1 MAAAEPSRRVGVGAGRNAAGLIRAGKVEAQHILASAPTDRLNLCFQALGRTH 60
|||||
Qy 61 SNOEVLQSCLLVIFATKPHVLPVLAELVSVAAAGVSLSTLEELLPPNTRY 120

```

Db      61  SNOEVLSQCLLVIFATKPHVLPVLAFAVAVVTTTEHILVSAAGVSLSTLEELLPPNTRV 120
QY      121  LRVLPNLPVVOEGALVMARGHRVGSSETKLLOHLLLEACGRCEEVPEAYVDIHTGLSSG 180
Db      121  LRVLPNLPVVOEGALVMARGHRVGSSETKLLOHLLLEACGRCEEVPEAYVDIHTGLSSG 180
QY      181  VAFVCAFSALAEAGAVKMGMPSSLAHRIAATQTLTGAKMLLHEGQHPAQLRSDVCTPGGT 240
Db      181  VAFVCAFSALAEAGAVKMGMPSSLAHRIAATQTLTGAKMLLHEGQHPAQLRSDVCTPGGT 240
QY      241  TIYGLHALEOGLRAATMSAVEAATCRAKELSRK 274
Db      241  TIYGLHALEOGLRAATMSAVEAATCRAKELSRK 274

RESULT 2
Q96HX4 PRELIMINARY: PRT: 274 AA.
AC Q96HX4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007993; AAH07993.1; -.
DR InterPro; IPR000304; P5CR.
DR Pfam; PF01089; P5CR; 1.
DR TIGRFAMs; TIGR00112; proc; 1.
DR PROSITE; PS00521; P5CR; 1.
KW Hypothetical protein.
SQ SEQUENCE 274 AA; 28649 MW; 846FD9B60183B048 CRC64;

```

```

Query Match 99.6%; Score 1381; DB 4; Length 274;
Best Local Similarity 99.6%; Pred. No. 9.7e-103;
Matches 273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAAEPSRRVGVGAGRMAGAIAGQLIRAGKVEAQHILASAPTDRLNLCHEFQALGCRTH 60
Db 1 MAAAEPSRRVGVGAGRMAGAIAGQLIRAGKVEAQHILASAPTDRLNLCHEFQALGCRTH 60
QY 61 SNOEVLSQCLLVIFATKPHVLPVLAFAVAVVTTTEHILVSAAGVSLSTLEELLPPNTRV 120
Db 61 SNOEVLSQCLLVIFATKPHVLPVLAFAVAVVTTTEHILVSAAGVSLSTLEELLPPNTRV 120
QY 121 LRVLPNLPVVOEGALVMARGHRVGSSETKLLOHLLLEACGRCEEVPEAYVDIHTGLSSG 180
Db 121 LRVLPNLPVVOEGALVMARGHRVGSSETKLLOHLLLEACGRCEEVPEAYVDIHTGLSSG 180
QY 181 VAFVCAFSALAEAGAVKMGMPSSLAHRIAATQTLTGAKMLLHEGQHPAQLRSDVCTPGGT 240
Db 181 VAFVCAFSALAEAGAVKMGMPSSLAHRIAATQTLTGAKMLLHEGQHPAQLRSDVCTPGGT 240
QY 241 TIYGLHALEOGLRAATMSAVEAATCRAKELSRK 274
Db 241 TIYGLHALEOGLRAATMSAVEAATCRAKELSRK 274

RESULT 3
Q8N3N9 PRELIMINARY: PRT: 244 AA.
AC Q8N3N9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

```

```

DE Hypothetical protein (Fragment).
GN DFEZP761H0716.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygala;
RA Ausorge W., Wirkner U., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL833857; CAD38716.1; -.
DR InterPro; IPR000304; P5CR.
DR Pfam; PF01089; P5CR; 1.
DR TIGRFAMs; TIGR00112; proc; 1.
DR PROSITE; PS00521; P5CR; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 244 AA; 25684 MW; 421656FB0FBA0741 CRC64;

Query Match 89.1%; Score 1235; DB 4; Length 244;
Best Local Similarity 99.6%; Pred. No. 4.1e-91;
Matches 243; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 31 GKVEAQHILASAPTDRLNLCHEFQALGCRTHSNOEVLSQCLLVIFATKPHVLPVLAFAVAP 90
Db 1 GKVEAQHILASAPTDRLNLCHEFQALGCRTHSNOEVLSQCLLVIFATKPHVLPVLAFAVAP 90
QY 91 VVTEHILVSAAGVSLSTLEELLPPNTRVLRVLPNLPVVOEGALVMARGHRVGSSETK 150
Db 61 VVTEHILVSAAGVSLSTLEELLPPNTRVLRVLPNLPVVOEGALVMARGHRVGSSETK 120
QY 151 LLOHLLLEACGRCEEVPEAYVDIHTGLSGVAFVCAFSALAEAGAVKMGMPSSLAHRIA 210
Db 121 LLOHLLLEACGRCEEVPEAYVDIHTGLSGVAFVCAFSALAEAGAVKMGMPSSLAHRIA 180
QY 211 QTLTGAKMLLHEGQHPAQLRSDVCTPGGTIIYGLHALEOGLRAATMSAVEAATCRAKE 270
Db 181 QTLTGAKMLLHEGQHPAQLRSDVCTPGGTIIYGLHALEOGLRAATMSAVEAATCRAKE 240
QY 271 LSRK 274
Db 241 LSRK 244

RESULT 4
Q9DCC4 PRELIMINARY: PRT: 274 AA.
AC Q9DCC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 1110058B13RIK protein.
GN 1110058B13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 24 Seconds
(without alignments)

536.888 Million cell updates/sec

Title: US-09-806-536A-14

Perfect score: 1386

Sequence: 1 MAAAPSPRRVGVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	584.5	42.2	1 PROC_HUMAN	P32322 homo sapien
2	533.5	38.5	1 PROC_METAC	Q9h999 methanosarc
3	502.5	36.3	1 PROC_PEA	Q04708 pisum sativ
4	498	35.9	1 PROC_ACTCH	Q04016 actinidia c
5	492	35.5	1 PROC_SOYBN	P17817 glycine max
6	475	34.3	1 PROC_ARATH	P54904 arabidopsis
7	423.5	30.6	1 PROC_ECOLI	P00373 escherichia
8	412.5	29.8	1 PROC_STNY3	P74572 synecocyst
9	390.5	28.2	1 PROC_TREPA	P27771 treponema p
10	387	27.9	1 PROC_THETH	P54893 thermus the
11	383.5	27.7	1 PROC_MYCLE	P46725 mycobacteri
12	378	27.3	1 PROC_BACSU	P14383 bacillus su
13	366	26.4	1 PROC_MYCTU	Q11141 mycobacteri
14	362.5	26.2	1 PROC_CORGL	P45450 corynebacte
15	362	26.1	1 PROC_CAEEL	Q20848 caenorhabdi
16	345.5	24.9	1 PROC_PSEAE	P22008 pseudomonas
17	344	24.8	1 PROC_SCHPO	Q9p7y7 schizosacch
18	337.5	24.4	1 PROI_BACSU	P54552 bacillus su
19	321	23.2	1 PROC_YEAST	P32263 saccharomyc
20	318.5	23.0	1 PROC_AQUAE	O66553 aquifex ae
21	311.5	22.5	1 PROC_FASMO	Q9cpe8 pasteurella
22	310.5	22.4	1 PROC_NEUCR	P12641 neurospora
23	307	22.2	1 PROC_VIBAL	P52053 vibrio algi
24	305.5	22.0	1 PROC_ZALAR	Q12740 zallerion ar
25	301.5	21.8	1 PROC_HAEIN	P43869 haemophilus
26	198	14.3	1 PROC_BACSU	Q00777 bacillus su
27	189.5	13.7	1 CME4_BACSU	P39696 bacillus su
28	164	11.8	1 PROC_HELPJ	Q9ZK56 helicobacte
29	159	11.5	1 PROC_HELPY	P22573 helicobacte
30	125.5	9.1	1 PROC_METSM	P23523 methanobrev
31	114	8.2	1 GARR_ECOLI	P23523 escherichia
32	99.5	7.2	1 ALR_MYCAV	Q91888 mycobacteri
33	98.5	7.1	1 ACK1_VIBVU	Q8dah8 vibrio vuln

ALIGNMENTS

SEQUENCE COMPARISON 'A'

RESULT 1

ID	PROC_HUMAN	STANDARD;	PRT;	319 AA.
AC	P32322: Q96DI6;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Pyroline-5-carboxylate reductase (EC 1.5.1.2) (p5CR) (p5C reductase).			
GN	PYCRI.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=92112821; PubMed=1730675;			
RX	Dougherty K.M., Brandriss M.C., Valle D.;			
RA	"Cloning human pyroline-5-carboxylate reductase cDNA by			
RT	complementation in Saccharomyces cerevisiae."			
RL	J. Biol. Chem. 267:871-875(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=Placenta;			
RC	MEDLINE=22388257; PubMed=12477932;			
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,			
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences".			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) -> L-pyrroline-5-			
CC	carboxylate + NAD(P)H.			
CC	-!- PATHWAY: Proline biosynthesis; third (last) step.			
CC	-!- SUBUNIT: HOMODECAMER OR HOMODECAMER.			
CC	-!- SIMILARITY: BELONGS TO THE PYROLINE-5-CARBOXYLATE REDUCTASE			
CC	FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL: M77836; AAA36407.1; -
 CC EMBL: BC001504; AAH01504.1; -
 CC PIR: A41770; A41770.1; -
 CC Genbank: HGNC:19721; PYCRL.
 CC GK: P32322; -
 CC MIM: 179035; -
 CC GO: GO:0004735; F:pyrroline 5-carboxylate reductase activity; TAS.
 CC GO: GO:0006561; P:proline biosynthesis; TAS.
 CC InterPro: IPR000304; P5CR.
 CC Pfam: PF01089; P5CR; 1
 CC TIGRFAMs: TIGR00112; proc; 1.
 CC PROSITE: PS00521; P5CR; 1.
 CC Oxidoreductase; Proline biosynthesis; NADP.
 KW CONFLICT 155 T -> S (IN REF. 2).
 FT CONFLICT 155 T -> S (IN REF. 2).
 SQ SEQUENCE 319 AA; 33374 MW; F5E74B5BDFB475EF CRC64;

Query Match 42.2%; Score 584.5; DB 1; Length 319;
 Best Local Similarity 46.4%; Pred No. 2.3e-37;
 Matches 124; Conservative 47; Mismatches 89; Indels 7; Gaps 3;
 QY 11 VGVFGAGRMAGIAAGLIRAGKVEAQHILASAPTRDNLCHPQAL---GCRTHHSNQEVLQ 67
 DB 3 VGFAGAGLAFALAKGFTAGAGVLAHAKIMASSP-DMDLATVSAIRKMGVRLTPHNETVQ 61
 QY 68 SCLLVIFATKPHVLPVAVLAPVAVVTEHILVSVAAAGVSLSTLEELP---PNTRVLRLV 124
 DB 62 HSDVFLAVKPHLPIIFLDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPRVLRM 121
 QY 125 PNLPCVQAGIYVARGHVGSSYKLLQHLLEACGRCEVEPEAYVDIHTGLSGSVAFV 184
 DB 122 TNPVVRGAVYATGTGHAQVEDGRLEQMLSTVGFTVEEDLIDAVTGLSGSPAYV 181
 QY 185 CAPSEALAGAVKMGMPSSLAHRIAATLLCTAKMLLHGGHQAOLRSDVCTPGGTIYG 244
 DB 182 FTALDALDGGVKGMLPRELAVLGAQLGAAKMLLHSEHPQLKDNVSSPGGATIH 241
 QY 245 LHLEGGRLAATMSAVEAATCRAKEL 271
 DB 242 LHVLEGGFRSLINAVEASCIRTREL 268

RESULT 2
 PROC_METAC STANDARD; PRT; 270 AA.
 AC Q9HH99; -
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
 GN PROC OR MA4102
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21833514; PubMed=11844777;
 RA Zhang J.K., White A.K., Kuetner H.C., Boccazzi P., Metcalf W.W.;
 RT "Directed mutagenesis and plasmid-based complementation in the
 RT methanogenic archaeon Methanosarcina acetivorans C2A demonstrated by
 RT genetic analysis of proline biosynthesis."
 RA J. Bacteriol. 184:1449-1454 (2002).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA an J.E., Nussbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RT W., Calvo S., Smirnov S., Smirnov S., Atnoor D., Brown A.,
 RT Naylor J., Stange-Thomann N., DeAtellano K., Johnson R.,

RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jiang H., Macario A.J.L., Paulsen I.,
 RA Perrychett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity";
 RL Genome Res. 12:532-542 (2002).
 CC -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) - 1-pyrroline-5-
 CC carboxylate + NAD(P)H.
 CC -!- PATHWAY: Proline biosynthesis; third (last) step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: AF305580; AAC22033.1; -
 DR EMBL: AE011122; AAM07450.1; -
 DR InterPro: IPR000304; P5CR.
 DR Pfam: PF01089; P5CR; 1.
 DR TIGRFAMs: TIGR00112; proc; 1.
 DR PROSITE: PS00521; P5CR; 1.
 KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 SQ SEQUENCE 270 AA; 27943 MW; 50EC656AFCL0B1CF CRC64;

Query Match 38.58; Score 533.5; DB 1; Length 270;
 Best Local Similarity 43.48; Pred. No. 1.4e-33;
 Matches 116; Conservative 53; Mismatches 97; Indels 1; Gaps 1;
 QY 9 RRVFGVAGRMAGIAAGLIRAGKVEAQHILASAPTRDNLCHPQAL---GCRTHHSNQEVLQ 67
 DB 4 QKIGFIGAGKMGSAALMQGTIKAGITVTPENIGASDVVEPFLKDLQAKLGRVSTDNVAVR 63
 QY 68 SCLLVIFATKPHVLPVAVLAPVAVVTEHILVSVAAAGVSLSTLEELP---PNTRVLRLV 127
 DB 64 ESDILILAVPQTLSSVLSNKKNEITSEKLVISIAAGVPLSTYEDALLGTRVVRVMPNI 123
 QY 128 PCVVQEGAIYVARGHVGSSYKLLQHLLEACGRCEVEPEAYVDIHTGLSGSVAFVCAF 187
 DB 124 AATVSEAAAGIAPGKNATPEDLKAALIFSAVGTAVQVPESLMDAVTGLSGSPAFIPV 183
 QY 188 SEALAEGAVKMGMPSSLAHRIAATLLCTAKMLLHGGHQAOLRSDVCTPGGTIYGLHA 247
 DB 194 TEAWADGAVLEGMDRKSAITLAAQTVLGAAKMALETGMHPGKDKVTSPTAGTTIQQHHS 243
 QY 248 LEOGGLRAATMSAVEAATCRAKELSRK 274
 DB 244 LEEAGIRAAFMNAVIRASERSKELGKK 270

RESULT 3
 PROC_PEA STANDARD; PRT; 273 AA.
 ID PROC_PEA
 AC Q04708;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
 GN PROC.
 OS Pisum sativum (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 24 Seconds
(without alignments)

536.888 Million cell updates/sec

Title: US-09-806-536A-14

Perfect score: 1386

Sequence: 1 MAAEPPRRVGFVGGRMA.....AATMSAVEATCRAKELSRK 274

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	584.5	42.2	319	1 PROC_HUMAN	P32322 homo sapien
2	533.5	38.5	270	1 PROC_METAC	Q9hh99 methanosarc
3	502.5	36.3	273	1 PROC_PEA	Q04708 pisum sativ
4	498	35.9	278	1 PROC_ACTCH	Q04016 actinidia c
5	492	35.5	274	1 PROC_SOYBN	P17817 glycine max
6	475	34.3	276	1 PROC_ARATH	P54904 arabidopsis
7	423.5	30.6	269	1 PROC_ECOLI	P00373 escherichia
8	412.5	29.8	267	1 PROC_SYNY3	P74572 synechocyst
9	390.5	28.2	263	1 PROC_TREPA	P27771 treponema p
10	387	27.9	261	1 PROC_TRETH	P54893 thermus the
11	383.5	27.7	294	1 PROC_MYCLE	P46725 mycobacteri
12	378	27.3	297	1 PROC_BACSU	P14383 bacillus su
13	366	26.4	295	1 PROC_MYCTU	O11141 mycobacteri
14	362.5	26.2	270	1 PROC_CORGL	P46540 corynebacte
15	362	26.1	299	1 PROC_CAEEL	Q20848 caenorhabdi
16	345.5	24.9	272	1 PROC_PSEAE	P22008 pseudomonas
17	344	24.8	282	1 PROC_SCHPO	Q9p7y7 schizosacch
18	337.5	24.4	278	1 PROI_BACSU	P54552 bacillus su
19	321	23.2	286	1 PROC_YEAST	P32263 saccharomyc
20	318.5	23.0	265	1 PROC_AQUAE	O66553 aquifex aeo
21	311.5	22.5	275	1 PROC_FASMU	Q9cpe8 pasteurella
22	310.5	22.4	311	1 PROC_NEUCR	O12641 neurospora
23	307	22.2	278	1 PROC_VIBAL	P52053 vibrio algi
24	305.5	22.0	320	1 PROC_ZALAR	Q12740 zallerion ar
25	301.5	21.8	271	1 PROC_HAEIN	P43869 haemophilus
26	198	14.3	272	1 PROC_BACSU	O00777 bacillus su
27	189.5	13.7	273	1 CME4_BACSU	P39696 bacillus su
28	164	11.8	257	1 PROC_HELPU	Q92K56 helicobacte
29	159	11.5	257	1 PROC_HELPU	O25773 helicobacte
30	125.5	9.1	251	1 PROC_METSM	P22350 methanobrev
31	114	8.2	294	1 GARR_ECOLI	P23523 escherichia
32	99.5	7.2	388	1 ALR_MYCAV	Q91888 mycobacteri
33	98.5	7.1	319	1 ACK1_VIBVO	Q8da88 vibrio vuln

RESULT 1

ID	PROC_HUMAN	STANDARD;	PRT;	319 AA.
AC	P32322: Q96DI6;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (p5CR) (p5C reductase).			
GN	PYCRI.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCHI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92112821; PubMed=1730675;			
RT	Dougherty K.M., Brandriss M.C., Valle D.;			
RT	"Cloning human pyrroline-5-carboxylate reductase cDNA by			
RT	complementation in Saccharomyces cerevisiae."			
RL	J. Biol. Chem. 267:871-875(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klauser R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = l-pyrroline-5-			
CC	carboxylate + NAD(P)H.			
CC	-!- PATHWAY: Proline biosynthesis; third (last) step.			
CC	-!- SUBUNIT: HOMODECAMER OR HOMODECAMER.			
CC	-!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE			
CC	FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			

O95255 homo sapien
Q9hxx3 pseudomonas
O53243 mycobacteri
P38056 mycobacteri
P98160 homo sapien
Q9cbr9 mycobacteri
Q87mz4 vibrio para
O33116 mycobacteri
Q9bzc7 homo sapien
Q9rvgl deinococcus
O88202 rattus norv
O50705 mycobacteri

ALIGNMENTS


```

CC -!- PATHWAY: Proline biosynthesis; third (last) step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES, BUT MOSTLY IN
CC NODULES.
CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16352; CAA34401.1; -
CC PIR; S10186; S10186.
CC InterPro; IPR000304; P5CR.
CC Pfam; PF01089; P5CR; 1.
CC TIGRfams; TIGR00112; proc; 1.
CC PROSITE; PS00521; P5CR; 1.
CC Oxidoreductase; Proline biosynthesis; NADP.
CC SEQUENCE 274 AA; 28586 MW; 933CFDFDD7598B63 CRC64;
CC -----
CC Query Match 35.5%; Score 492; DB 1; Length 274;
CC Best Local Similarity 39.6%; Pred. No. 2e-30;
CC Matches 107; Conservative 51; Mismatches 96; Indels 16; Gaps 3;
CC -----
CC QY 11 VGFVAGRMAGATAOGLIRAGKVEAQHILASAPTRNLCH-----FOALGCRTHSN 62
CC Db :||:||||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
CC 13 LGFVAGKMAESITARGVSGVLPSPRI-----RTAVHFNLAARRGAFESFGVTLPSPN 65
CC QY 63 QEVLOSCLLVIFATKPHVLPVLAELVAPVVTTHILSVSAAGVSLTLEELLPPNTRVLR 122
CC Db :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
CC 66 DDVVRESDVVLSVKPQLKDVVSKLTPLLTTHKLLVSAAGTKLKLQEQ-WAGNDREIR 124
CC QY 123 VLPNLPVVOEGAIVMARGHRVGSSETKLLQHLLEACGCEVPEAYVDIHTGLSGGVA 182
CC Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 125 VNPNTAAVQQAASVNSGLSGSATEEDGNTIIAFGSIKWKAEKDYFDATITGLSGGPA 184
CC QY 183 FVCAFSEALAEAGVAKMGMPSSLAHRIAQTLLGTAKMLLHGHQHPAQLRSDVCTPGGTTI 242
CC Db :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
CC 185 YVYVLAELADGVAAGLPDRDLSLSASQTLVLSAASVMSQTKHPGQLKDDVTSFGGTTI 244
CC QY 243 YGLHALEOGLRAAATNSAVEATCRAKELS 272
CC Db :||:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 245 TGIHELENGFRGRTLMNAVVAARSRRELS 274
CC -----
CC RESULT 6
CC PROC_ARATH STANDARD; PRT: 276 AA.
CC AC P54904;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Pyrraline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
CC GN PROCI OR AT5G14800 OR T9L3.100.
CC OS Arabidopsis thaliana (Mouse-ear cress).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
CC OX NCBI_TaxID=3702;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=94294559; PubMed=8022935;
CC RA Verbruggen N., Villarroel R., van Montagu M.;
CC RT "Osmoregulation of a pyrraline-5-carboxylate reductase gene in
CC Arabidopsis thaliana."
CC RL Plant Physiol. 103:771-781(1993).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=cv. Landsberg erecta;

```

```

RA Verbruggen N., Villarroel R., Hua X., van Montagu M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RX MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Makayama S.,
RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yasuda M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latrille P.,
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Maria M.,
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA Ransperger U., Wedler H., Baïke K., Wedler E., Peters S.,
RA Van Staveren M., Dirkse W., Moolijman P., Klein Lankhorst R.,
RA Weitzenecker T., Bothe G., Rose M., Hauf J., Berneiser S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT thaliana."
RL Nature 408:823-826(2000).
CC -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC carboxylate + NAD(P)H.
CC -!- PATHWAY: Proline biosynthesis; third (last) step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M76538; AAA61346.1; -
CC EMBL; Y08951; CAA70148.1; -
CC EMBL; AL391149; CAC01879.1; -
CC PIR; JQ2334; JQ2334.
CC InterPro; IPR000304; P5CR.
CC Pfam; PF01089; P5CR; 1.
CC TIGRfams; TIGR00112; proc; 1.
CC PROSITE; PS00521; P5CR; 1.
CC KW Oxidoreductase; Proline biosynthesis; NADP.
CC SEQUENCE 276 AA; 28624 MW; B577A01C92A3A28B CRC64;
CC -----
CC Query Match 34.3%; Score 475; DB 1; Length 276;
CC Best Local Similarity 40.0%; Pred. No. 3.8e-29;
CC Matches 106; Conservative 54; Mismatches 103; Indels 2; Gaps 2;
CC -----
CC QY 10 RVGFVAGRMAGATAOGLIRAGKVEAQHILASAPTRNLCH-FOALGCRTHSNQVLOS 68
CC Db :||:||||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:| :||:|
CC 12 KVGFVAGKMAESITARGVSGVLPSPRICTAVHFNLAARRDVFESGVNFTSEEVKKE 71
CC QY 69 CLLVIFATKPHVLPVLAELVAPVVTTHILSVSAAGVSLTLEELLPPNTRVLRVLPNLP 128
CC Db :||:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 72 SDVVIFSVKPVVK&AVTELKSKLSKNKILSVSAAGIKLNDLQEQ-WSGODRFRVMPNTP 130
CC QY 129 CVVQEGAIVMARGHRVGSSETKLLQHLLEACGCEVPEAYVDIHTGLSGGVAFCAPS 188
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 131 AAVGEAASVSLGTGATEDEGAIVAMLFAGVCKILKADEKMFDAVTGLSGSGPAXIFIAI 190
 QY 189 EALAEGAVKMGPSLSLAHRIAQAOTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHAL 248
 DDb 191 EALADGGVAAGLAPRELALSLSQTVLGAATMYSGTKGHPVLKDDVTSFGGTTIAGVHEL 250
 QY 249 EQGGLRAATYMSAVEAATCRAKELSR 273
 DDb 251 EKGFSFRATLNAVVAAKRSRELSQ 275

RESULT 7

PROC_ECOLI STANDARD: PRT; 269 AA.
 AC P00373;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
 GN PROC OR B0386 OR SF0322.
 OS Escherichia coli, and
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 623;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC SPECIES=E.coli;
 RX MEDLINE=83116986; PubMed=6296787;
 RA Deutch A.H., Smith C.J., Rushlow K.E., Kretschmer P.J.;
 RT "Escherichia coli delta 1-pyroline-5-carboxylate reductase: gene
 RT sequence, protein overproduction and purification."
 RL Nucleic Acids Res. 10:7701-7714(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli;
 RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
 RA Pederspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H.,
 RA Lin D., Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=2272406; PubMed=12384590;
 RA Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157."
 RL Nucleic Acids Res. 30:4432-4444(2002).
 CC -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyroline-5-
 CC carboxylate + NAD(P)H.
 CC -!- PATHWAY: Proline biosynthesis; third (last) step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE PYROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL; J01665; AA86433.1; -;
 DR EMBL; AE000145; AAC7349.1; -;
 DR EMBL; U73857; AAB1810.1; -;
 DR EMBL; AE015066; AAN41981.1; -;
 DR PIR; A00385; RDECC.
 DR EcoGene; EG10769; proc.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF01089; P5CR; 1.
 DR TIGRFAMS; TIGR00112; proc; 1.
 DR PROSITE; PS00521; P5CR; 1.
 KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 SQ SEQUENCE 269 AA; 28145 MW; C5E2CED368CC508D CRC64;

Query Match 30.6%; Score 423.5; DB 1; Length 269;
 Best Local Similarity 34.6%; Pred. No. 3e-25;
 Matches 92; Conservative 56; Mismatches 117; Indels 1; Gaps 1;

QY 9 RRVGFVGAGRMAGATAOGLIRAGKVEAOHILASAPT-DRNLCHFQALGCRTHTHSNQEVILQ 67
 DDb 3 KKIGFIGCGNMKAILGLIASGOVLPQGIWVYTPSPDKVAALHDQFGINAAESAQAEVQAQ 62
 QY 68 SCLLVIFATKPHVLPVLAELAVPVVVTTEHILVSVAAGVSLSTLELLPPNTRVLRLVLPNL 127
 DDb 63 IADIIFRAVKPGIMTKVLTSSLNKDSLVSIAAGVTLDLQARALGHDRKIIRAMPNT 122
 QY 128 PCVVOEGAVMARGHRVGSSTKLLQHLLEACGRCREEVPEAYVDIHTGLSGSGVAFVCAF 187
 DDb 123 PALVNAAGTSTVTPNALVTPEDADVNLNIFRCFGEAEVIAEPMIHPVVGVSGSSPAYVMEF 182
 QY 188 SEALAEAGVAKMGPSSLAHRIAQAOTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
 DDb 183 TEAMADAALVIGMPRAQYFAAQAVMGSAKMVLETGEHPGALKDMVCSPGTTIEAVRY 242
 QY 248 LEQGLRAATMSAVEAATCRAKELSR 273
 DDb 243 LEKGFRAAVTEAMTKMEKSEKLSK 268

RESULT 8

PROC_SYNY3 STANDARD: PRT; 267 AA.
 AC P74572;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
 GN PROC OR SLR0661.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hikosawa M., Sugliura M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-136(1996).
 CC -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrolone-5-
 CC carboxylate + NAD(P)H.
 CC -!- PATHWAY: Proline biosynthesis; third (last) step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE PYROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.


```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pyroline-5-carboxylate reductase 1 (EC 1.5.1.2) (P5CR 1) (P5C
DE reductase 1).
GN PROH OR PROC.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessi eres P., Polotini A., Borchert S.,
RA Borri s R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.M., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.J., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goireau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzm tger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT *The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.*;
RL Nature 390:249-256(1997).
RL [2]
RP SEQUENCE OF 1-256 FROM N.A.
RC STRAIN=W23, and 168;
RX MEDLINE=91192601; PubMed=1849493;
RA Ann K.S., Wake R.G.;
RT *Variations and coding features of the sequence spanning the
RT replication terminus of Bacillus subtilis 168 and W23 chromosomes.*;
RL Gene 98:107-112(1991).
RL [3]
RP SEQUENCE OF 1-200 FROM N.A.
RC STRAIN=W23;
RX MEDLINE=89155440; PubMed=2493444;
RA Lewis P.J., Wake R.G.;
RT *DNA and protein sequence conservation at the replication terminus in
RT Bacillus subtilis 168 and W23.*;
RL J. Bacteriol. 171:1402-1408(1989).
RL [4]
RP SEQUENCE OF 199-297 FROM N.A., AND FUNCTION.
RC STRAIN=168 / SMY;
RX MEDLINE=21311767; PubMed=11418582;
RA Belitsky B.R., Brill J., Bremer E., Sonenshein A.L.;
RT *Multiple genes for the last step of proline biosynthesis in Bacillus
RT subtilis.*;
RL J. Bacteriol. 183:4389-4392(2001).
CC -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+)=1-pyrroline-5-
CC carboxylate + NAD(P)H.
CC -!- PATHWAY: Proline biosynthesis; third (last) step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE PYROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99114; CAB13741.1; --
CC EMBL; M24523; AAA22722.1; --
CC EMBL; AF006720; AAB62697.1; --
CC PIR; E69682; E69682.
CC Subtilist; BG11049; proH.
CC InterPro; IPR000304; P5CR.
CC Pfam; PF01089; P5CR; 1.
CC TIGRFAMs; TIGR00112; proC; 1.
CC PROSITE; PS00521; P5CR; 1.
CC Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
KW VARIANT 58 58 T -> A (IN STRAIN W23).
FT VARIANT 70 70 L -> S (IN STRAIN W23).
FT VARIANT 74 74 L -> I (IN STRAIN W23).
FT VARIANT 93 93 N -> S (IN STRAIN W23).
FT VARIANT 100 100 S -> T (IN STRAIN W23).
FT VARIANT 127 127 E -> Q (IN STRAIN W23).
FT VARIANT 159 159 K -> Q (IN STRAIN W23).
FT VARIANT 236 236 H -> Q (IN STRAIN W23).
FT VARIANT 239 239 I -> V (IN STRAIN W23).
FT CONFLICT 256 256 Q -> E (IN REF. 2).
FT CONFLICT 264 297 GEALSOAKHAAKSKSEIDIEKTAAPLSGVK -> EKR
SQ SEQUENCE 297 AA; 32031 MW; F4DC7E6F84C6ED42 CRC64;
Query Match 27.3%; Score 378; DB 1; Length 297;
Best Local Similarity 32.2%; Pred. No. 9.2e-22;
Matches 91; Conservative 50; Mismatches 106; Indels 36; Gaps 3;
QY 9 RRVGVAGRMAGATAOGLIRAGKVEAOHILASAPTDRNLCHFQALGCTHHSNOEVL-- 66
Db 17 KKVAFIGASMAEGMISGIVRANKIPKQNI-----CVTVRSWTERLTE 59
QY 67 -----QSLC---LVIFATPKHVLPAVLAEPVVTTEHILVSVAAGVSLST 109
Db 60 LELOYGIKALPNQLCIEDMDVLILAMPKDAENALSLKSRIOPHQLLSVLGAGITTSF 119
QY 110 LEELPPNTRVLRLPNLPCVVQEGATVMARGRHVGSSETKLQLHLEACGRCEVPEAY 169
Db 120 IQSLLNEQPVVRVMPNTSSMIGASATAIALGKYVSEDLKKLAELLCMGVEVYTIQENQ 179
QY 170 VDIHTGLSGSVAFVCAFSEALAECAVKMGMPSSLAHRIAATQTLTGAKMLLHGGQHPAQ 229
Db 180 MDIFTGIAGSGPAYFYMEFIEKTEGEGAGLDKQLSRISIGAGTLLGAALKMLMETGEHPEI 239
QY 230 LRSDVCTPGGTTIYGLHLEOGLRAATMSAVEAATCAKELS 272
Db 240 LRDNITSPNGTTAAGLQALKKSGGGEATSOALKHAAKRSKETS 282
RESULT 13
PROC_MYCTU STANDARD; PRT: 295 AA.
AC Q11141;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pyroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN PROC OR RV0500 OR MT0520 OR MTCY2069.26.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RN SEQUENCE FROM N.A.

```


